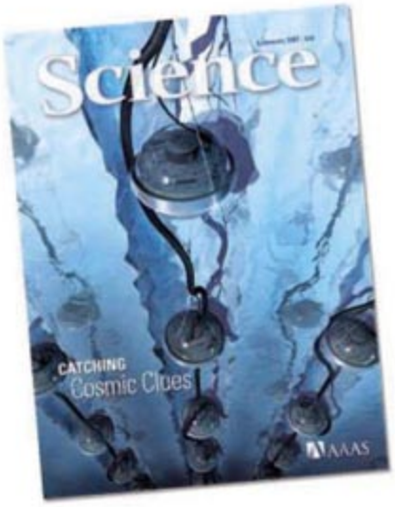
5 January 2007 | S10

Science

CATCHING
Cosmic Clues

AAAS

Volume 315, Issue 5808



COVER

Artist's conception of the IceCube array of photodetectors now under construction at the South Pole. When complete, IceCube will detect neutrinos originating from collisions of cosmic rays with nitrogen and oxygen in the Northern Hemisphere; neutrinos reaching the detector must first pass through the entire planet. See the special section on particle astrophysics beginning on page 55.

Image: Chris Bickel/Science

For related online content, see page 13

DEPARTMENTS

EDITORIAL

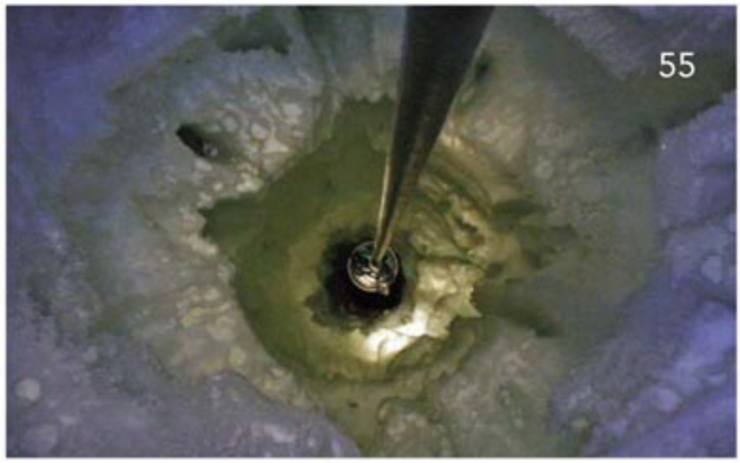
SPECIAL SECTION

Particle Astrophysics

INTRODUCTION

NEWS

PERSPECTIVES



NEWS OF THE WEEK

NEWS FOCUS

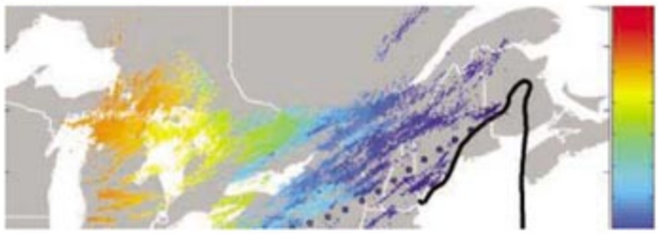
CONTENTS continued >>

SCIENCE EXPRESS

www.sciencexpress.org

ATMOSPHERIC SCIENCE

Direct Measurements of the Convective Recycling of the Upper Troposphere
T. H. Bertram et al.
Direct measurements imply that the rate of deep convection in the troposphere may be faster than predicted, affecting our understanding of chemical reactions in air.
10.1126/science.1134548



APPLIED PHYSICS

Coding/Decoding and Reversibility of Droplet Trains in Microfluidic Networks
M. J. Fuerstman, P. Garstecki, G. M. Whitesides
A microfluidics device in which droplets move into one of two unequal channels shows reversible nonlinear effects that can be manipulated to encode and decode signals.
10.1126/science.1134514

MEDICINE

An X Chromosome Gene, *WTX*, Is Commonly Inactivated in Wilms Tumor
M. N. Rivera et al.
The identification of a gene mutated in pediatric kidney cancer suggests that genes located on the X chromosome play a greater role in cancer than has been thought.
10.1126/science.1137509

LETTERS

BOOKS *ET AL.*

POLICY FORUM

PERSPECTIVES

BREVIA

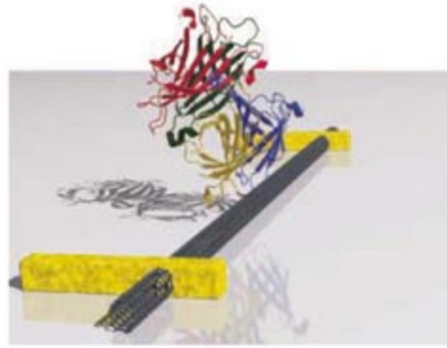
GENETICS

REPORTS

PHYSICS

CHEMISTRY

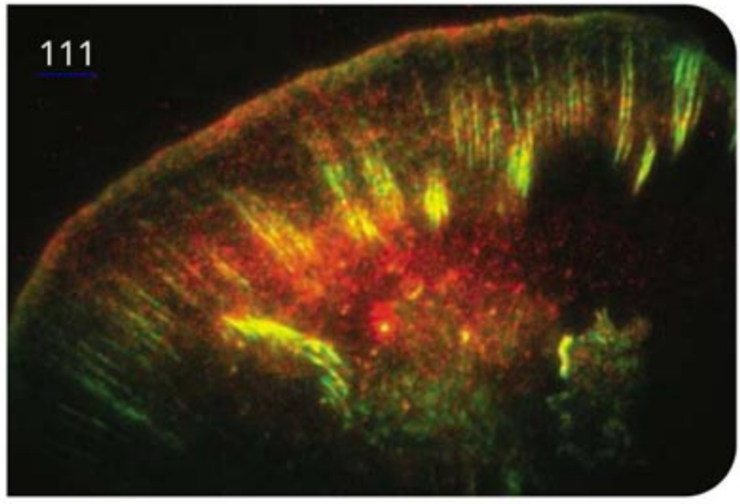
CHEMISTRY



77

CONTENTS continued >>

REPORTS *CONTINUED...*



AAAS
ADVANCING SCIENCE, SERVING SOCIETY

SCIENCE (ISSN 0036-8075) is published weekly on Friday, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Periodicals Mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices. Domestic individual membership and subscription (51 issues): $142 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $710; Foreign postage extra: Mexico, Caribbean (surface mail) $55; other countries (air assist delivery) $85. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #1254 88122. Publications Mail Agreement Number 1069624. **Printed in the U.S.A.**

Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. **Postmaster:** Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090–6178. **Single-copy sales:** $10.00 current issue, $15.00 back issue prepaid includes surface postage; bulk rates on request. The identification code for *Science* is 0036-8075. *Science* is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

CONTENTS continued >>

SCIENCENOW

www.sciencenow.org DAILY NEWS COVERAGE

'Social Dementia' Decimates Special Neurons
Study ties inappropriate behavior to loss of recently discovered spindle neurons.

Could Bioterror Warnings Make You Sick?
Media messages about biological attacks may have a downside, researchers say.

Delving Into Chloroplasts' Past
A tiny genome shows that green algae twice became slaves of other cells.

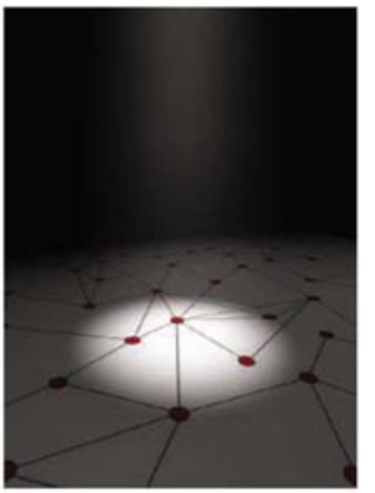
Are you ready for electronic grant applications?

SCIENCECAREERS

www.sciencecareers.org CAREER RESOURCES FOR SCIENTISTS

US: What You Need to Know About Electronic R01 Submissions
A. Kotok
The first deadline for electronic submission of NIH R01 grant proposals is approaching.

GRANTSNET: January 2007 Funding News
J. Fernandez
Get the latest index of research funding opportunities, scholarships, fellowships, and internships.



Spotlight on signaling.

SCIENCE'S STKE

www.stke.org SIGNAL TRANSDUCTION KNOWLEDGE ENVIRONMENT

EDITORIAL GUIDE: 2006—Signaling Breakthroughs of the Year
E. M. Adler, N. R. Gough, L. B. Ray
Scientists nominate the year's most exciting research related to signal transduction.

REVIEW: How Do MicroRNAs Regulate Gene Expression?
R. J. Jackson and N. Standart
Researchers are closing in on mechanisms that allow microRNAs to control stability and translation of mRNAs.

Opportunities for particle astrophysicists.

SPECIAL SECTION

Particle Astrophysics

SCIENCECAREERS

www.sciencecareers.org CAREER RESOURCES FOR SCIENTISTS

GLOBAL: Special Feature—Particle Astrophysics
J. Austin
Particle astrophysicists are bringing together particle physics and astrophysics to the benefit of both fields.

FRANCE: Getting Big on Astroparticles
E. Pain
French astroparticle physicist Guillaume Dubus manages to keep a foot grounded in four scientific worlds.

US: Catching Some Cosmic Rays
A. Fazekas
Postdoc Vasiliki Pavlidou applies ideas from particle physics to discover more about gamma rays.

Separate individual or institutional subscriptions to these products may be required for full-text access.

EDITED BY STELLA HURTLEY AND PHIL SZUROMI

Ancient Aerobics >>

The emergence of animals in the Late Proterozoic (about 580 million years ago) may have been aided by the oxygenation of Earth's atmosphere and oceans, but little evidence for such an environmental change has been reported. **Canfield *et al.*** (p. 92, published online 7 December; see the 8 December news story by **Kerr**) examined the distribution of iron in rocks in Newfoundland that represent deep-ocean deposits from the Late Proterozoic. Their data imply that the deep ocean became oxygenated immediately after the last major Proterozoic glaciation. This change immediately preceded the appearance of the first animal fossils in these rocks.

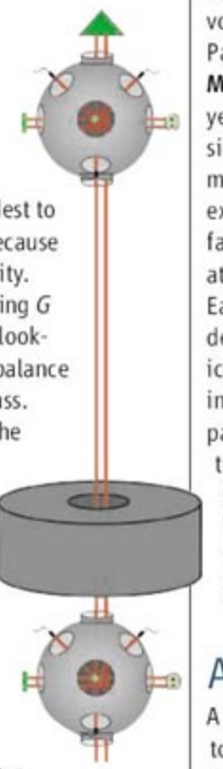
Sensitive Sidewalls

The conductivity of the sidewalls of single-walled carbon nanotubes (SWNTs) changes when they are modified by defects or adsorbed molecules and, like graphite, the sidewalls of bulk samples can be modified by electrochemical oxidation. **Goldsmith *et al.*** (p. 77) have exploited this sensitivity by performing electrochemical oxidation and reduction on individual SWNTs mounted on electrodes so that their conductance G can be monitored. Oxidation in strong acids caused stepwise drops in G that the authors attribute to the formation of C–O groups that bond to the acid's conjugate base. Reduction recovers most but not all of the drop in conductivity, indicating that rather than reforming the pristine sp^3 carbon framework, sp^2 groups with minimal electron scattering, such as ether linkages, can form instead.

Tracing a Stratospheric Journey

Large volcanic eruptions inject material into the stratosphere and impact global climate, but a lack of observational data has made it difficult to determine if an ancient volcanic eruption, which might only be documented by deposited ash layers, affected the stratosphere. **Baroni *et al.*** (p. 84) now report that the isotopic composition of the sulfur in sulfate contained in Antarctic snow for the Agung (1963) and Pinatubo (1991) eruptions displays mass-independent fractionation in the sulfate concentration peaks. Because only photochemical reactions in the stratosphere can explain this pattern of isotope fractionation, the authors suggest that sulfate isotopic composition could be used to record whether volcanic ash entered the stratosphere.

Massive Interference



The Newtonian gravitational constant G is the fundamental constant that has been the hardest to determine accurately, in part because of the relative weakness of gravity. Traditional methods for measuring G tend to be mechanical, such as looking at the rotation of a torsion balance in response to a moving test mass. **Fixler *et al.*** (p. 74) show that the interference pattern of the de Broglie waves of cold cesium atoms shifts in response to the position of a 540-kilogram test mass made of lead. The authors claim that the technique is less prone to the systematic errors that plague the mechanical measurements and may ultimately allow for a more accurate determination of G.

Greenhouse Gases in an Earlier Ice Age

Numerous studies of Cenozoic climate have shown how climate and the carbon cycle are linked, but similar records much farther back in time are rare. Before the start of the current "icehouse," around 35 million years ago (Ma) when large ice sheets began to form in Antarctica, the last period when Earth had sizable volumes of continental ice was during the late Paleozoic (between 265 and 305 Ma). **Montañez *et al.*** (p. 87) used a 40-million-year–long record of the stable isotopic compositions of minerals formed in soils, fossil plant matter, and shallow-water brachiopods to explore the relation between continental surface temperatures and the concentration of atmospheric CO_2 during this interval when Earth drifted in and out of glaciated and fully deglaciated conditions. Changes in continental ice volume were strongly correlated with shifts in atmospheric partial pressure of CO_2, and paleofloral data chronicle the repeated restructuring of paleotropical floral communities that accompanied the inferred climate shifts. These findings suggest that greenhouse gas forcing of climate occurred during remote times in a manner similar to the present era.

Altered Aerobics

A comparison of field observations and laboratory data on the eelpout (a North Sea and Baltic Sea fish species) by **Pörtner and Knust** (p. 95; see the Perspective by **Wang and Overgaard**) has revealed a mismatch between tissue oxygen supply and temperature-dependent oxygen demand that is causing a loss of species abundance during hot summers. Temperature-dependent constraints in oxygen supply are likely to affect many functions, includ-

CREDITS (TOP TO BOTTOM): GUY NARBONNE; FIXLER *ET AL.*

ing behavior, growth, reproduction, and interaction with other species, and thereby influence the long-term fate of populations and species in various climates.

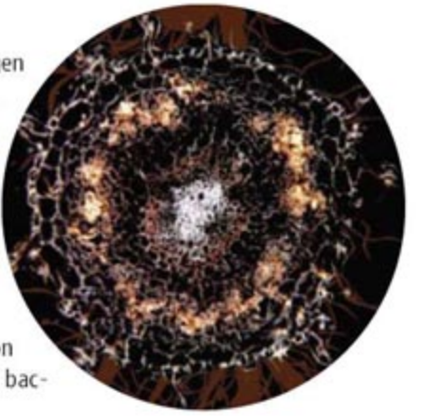
Targeting MicroRNAs to the Nucleus

MicroRNAs (miRNAs) are small ~22-nucleotide (nt) noncoding RNAs found in most eukaryotes that regulate the translation and/or stability of target RNAs. miRNAs are grouped into families that are related by their highly conserved 5′ "seed" sequences that are important in defining the complementary target RNAs. The 3′ sequences are generally less conserved within families, which has raised questions about their functional significance. Even so, 3′ sequences can be very highly conserved (even identical) across species for individual miRNAs, which suggests the presence of powerful selective constrains. **Hwang *et al.*** (p. 97) now show that human miR-29b is localized to the nucleus and that this localization is driven by a 6-nt sequence in the 3′ half of the molecule. The authors raise the intriguing possibility that miR-29b might regulate the transcription or splicing of target transcripts.

Biased Inheritance

Although chromosome segregation is generally considered to be random relative to daughter cell inheritance, nonrandom segregation of mouse chromosome 7 has been reported for certain cell types. **Armakolas and Klar** (p. 100; see the Perspective by **Sapienza**) examined molecular components that participate in nonrandom chromatid segregation. Mutation of a gene encoding the microtubule motor left-right dynein (LRD), shown previously to affect left-right body-axis determination, differentially affected chromatid segregation in specific cell types.

Giving Lotus the Nodule



The nodulation of roots in legumes is a key factor in nitrogen fixation. Working in the leguminous plant *Lotus japonica*, **Trichine *et al.*** (p. 104, published online 16 November) and **Murray *et al.*** (p. 101, published online 16 November) have identified how the hormone cytokinin fits into the signaling cascade by which leguminous plants establish nitrogen-fixation nodules filled with symbiotic bacteria (see the Perspective by **Oldroyd**). Gain-of-function mutation in a cytokinin receptor results in spontaneous formation of bacteria-free nodules, whereas loss of function results in too few nodules, despite aggressive formation of bacterial infection threads.

Choosing the Right Path

Dendritic cells of the immune system present antigen to T cells in the context of either class I or class II molecules of the major histocompatibility complex (MHC) as a means of generating two distinct arms of T cell immunity: Class I–restricted $CD8^+$ T cell responses and $CD4^+$ T cell help. **Dudziak *et al.*** (p. 107) present evidence that each pathway dominates in distinct subsets of dendritic cells. Using chimeric antibodies specific for cell surface markers present on each specific subtype of dendritic cell, it was possible to target antigens to the class I or class II MHC pathways and so elicit CD8 or CD4 responses, respectively.

Enzyme Kinetics in Living Color

Innovative methods are required to study the spatial regulation of enzymatic activity inside living cells. **Yudushkin *et al.*** (p. 115) describe such a method based on the detection of enzyme-substrate complexes using fluorescence lifetime imaging microscopy. The detection of the enzyme interacting with the substrate ensures utmost specificity and enables evaluation of the localized activity of a particular enzymatic species. The technique was used to investigate the spatial regulation of growth factor signaling by the tyrosine phosphatase PTP1B and revealed that PTP1B exists inside cells as kinetically distinct, spatially separated subpopulations.

CREDIT: K. SZCZYGLOWSKI AND A. MOLNAR/AGRICULTURE AND AGRI-FOOD CANADA

Donald Kennedy is the Editor-in-Chief of *Science*.

A New Year

AT THE BEGINNING OF EACH NEW YEAR, I HAVE DIFFICULTY DECIDING WHETHER TO FOCUS the Editorial on what's coming up, or on what's happened during the last one. This time around, there's just too much of both, so I have assumed the Mugwump posture to do a little of each.

Looking forward, an encouraging sign is that it was a good year for international science. A multinational experimental fusion reactor, ITER, got under way, and collaboration in space is growing more active. The new German head of state is a scientist, a Russian proved the Poincaré Conjecture, and Australia reversed its ban on stem cell research. As for the United States, midterm elections will deliver new occupants for important House chairmanships, and for the Senate too, if the slim majority holds. The Senate's Environment Committee gets Barbara Boxer of California, a huge contrast to incumbent James Inhofe of Oklahoma. Boxer finds the scientific evidence on climate change convincing, along with most of the rest of the country. Inhofe, on the other hand, is a conspiracy theorist who calls global warming a grand hoax. His farewell hearing, held as the 109th Congress limped off the stage, featured his usual crowd of skeptics.

Among the other new leaders is Bart Gordon for the House Science Committee (see News section). He replaces an excellent Republican Chairman, the newly retired Sherry Boehlert, so this switch is not a rescue but a quality succession. The Republican leadership helped out, ignoring seniority to name Ralph Hall ranking minority rather than former Chairman James Sensenbrenner. And the new House leadership announced a meeting schedule for 2007 that puts to shame the lackadaisical work habits of the 109th.

The bad news involves the budget. The new Congress inherits a Continuing Resolution (CR) that will last through February 15, locking spending at existing or even lower levels. That's bad enough, but recent word is that the incoming Democrats plan to extend the CR to the end of the fiscal year (FY) 2007 budget year at FY'06 levels. That would be a real loss for science funding. Their announced intention is to include financial fixes for programs that would be seriously hurt by the CR, but that's a pretty squishy promise, so we'll stay tuned to learn what will really happen to science budgets. Elsewhere, other governments are being more generous. Several European Union nations continue to grow their research expenditures, and China's allocation for 2005 was up by a stunning 16% from the preceding year.

Looking in our rearview mirror, which reminds us that objects often appear more distant then they really are, we find few sources for comfort and joy. It just wasn't a banner year. There was too much fraud in science, including a major case that put some egg on our face here. Little progress was seen in the United States and in Germany, for example, on stem cell research; and although intelligent design advocates lost in Pennsylvania and Kansas, the topic won't die. The holiday present from the U.S. Environmental Protection Agency (EPA) was another lump of coal in science's stocking. Here's how it will review the National Ambient Air Quality Standards for such "criteria pollutants" as particles, ozone, and soot. Instead of using outside science findings and EPA scientists, political appointees will collaborate with the latter to summarize "policy-relevant" science. The EPA's own Science Advisory Board doesn't like this much, presumably seeing it as yet another case of the policy cart leading the science horse.

But there are some real bright spots. New climate data, not to mention Al Gore's film and the determined position of the British government on this issue, seem to have altered the U.S. mood about global warming. Even Exxon Mobil, the most deeply embedded industrial holdout, is changing its tune; could Senator Inhofe be next? We close this topic with a fun item: here's U.S. Supreme Court Justice Antonin Scalia after a Massachusetts attorney had corrected his misuse of "stratosphere" in the carbon dioxide case: "Troposphere, whatever. I told you before I'm not a scientist. That's why I don't want to have to deal with global warming, to tell you the truth."

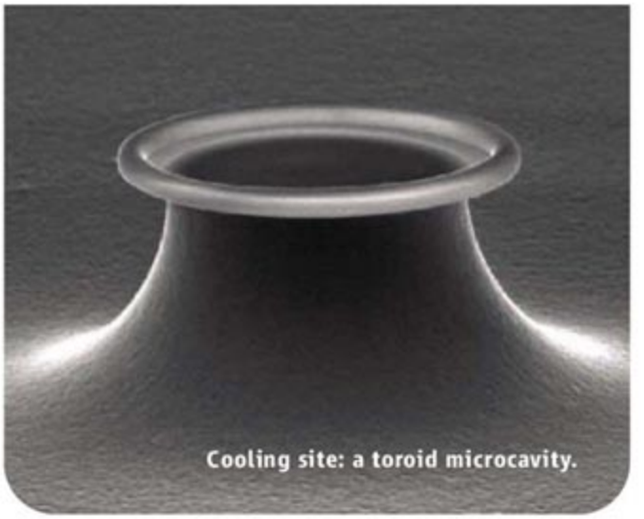
Troposphere, schmoposphere, your Honor. Happy New Year!

– Donald Kennedy

10.1126/science.1139394

APPLIED PHYSICS

Cooling Rays of Light


Cooling site: a toroid microcavity.

Just as the vibrational frequency of a mechanical oscillator shifts in response to changes in its environment (e.g., changes in the pressure, temperature, or viscosity of the medium in which it sits), so it may be expected that the radiation pressure exerted by light on an object can also affect the vibrational modes of mechanical resonators. This phenomenon opens the possibility of either amplifying (heating) or damping (cooling) the motion of the resonator with light. Whereas laser cooling is now routine for microscopic objects such as atoms, translating the technique to larger objects presents more of a challenge, because the dynamical back-action between the photons and the resonator requires that photon lifetimes be long enough to interact with the mechanical modes of the resonator. Effectively, the photons must be confined in the cavity on a time scale comparable to the mechanical oscillation period of the resonator. Four recent studies, by Schliesser *et al.*, Gigan *et al.*, Arcizet *et al.*, and Kleckner and Bouwmeester, successfully access this regime for dynamical back-action and demonstrate efficient optical cooling of a mechanical oscillator mode to cryogenic temperatures. The ability to cool macroscopic objects with light not only has practical applications, as for mirror stabilization in large-scale interferometers, but also offers a means of probing quantum effects in mechanical systems. — ISO

Phys. Rev. Lett. **97**, 243905 (2006); *Nature* **444**, 67; 71; 75 (2006).

ASTROCHEMISTRY

Capturing Ferroelectric Ice

At low temperature and pressure, water crystallizes in two distinct morphologies, termed ice I and ice XI. Ice I exhibits the form of a hexagonal lattice of oxygen atoms, with attached protons distributed randomly around them. In ice XI, the protons become ordered and the resulting solid is ferroelectric. The inherent stability of ice XI is of particular interest because of its possible formation in space. However, researchers have accessed it only by doping of water samples with potassium hydroxide, and the influence of the dopant on long-range ordering was not well resolved.

Fukazawa *et al.* have succeeded in making large quantities of Ice XI in the laboratory by doping D_2O (deuterated to raise the neutron scattering efficiency) with very small amounts of KOD, and then carefully maintaining the samples in a 60 to 70 K temperature range over tens of hours. Neutron diffraction experiments confirmed an extended ordered structure. The existence of ice XI in cold space environments is therefore likely; the electronic properties of the bulk ice may affect the formation mechanism of icy planets. — JB

Astrophys. J. **652**, L57 (2006).

GENETICS

Pining for Understanding

The genes underlying complex (and industrially important) traits in pine have long been sought, but the paucity of genetic resources has made this an arduous search. González-Martínez *et al.* use a population genomic approach to examine the associations between phenotypic traits and single-nucleotide polymorphisms (SNPs) in known genes to identify specific allelic variants underlying solid wood production and wood biochemistry in loblolly pine. In spite of the large genome size in conifers, the high heterozygosity and rapid breakdown of linkage disequilibrium allowed them to identify 20 genes underlying complex polymorphic traits. Although the effects demonstrated for each SNP were relatively low, on the order of 5% (similar to that observed in previously identified quantitative trait loci), combining markers associated with the same trait accounted for 20% of the phenotypic variation and 40% of the additive genetic variance.

Besides its potential commercial use in tree breeding, this approach can also be applied to investigations of the evolution and ecological genetics of loblolly pine. — LMZ

Genetics 10.1534/genetics.106.061127 (2006).

DEVELOPMENT

Import Controls

The directed and controlled differentiation of cells is of critical importance for being able to use embryonic stem cells in a clinical setting. Yasuhara *et al.* have shown that a switch in a nuclear transport mechanism is involved in cell fate determination. For nuclear import, a protein with a nuclear localization signal (NLS) binds to the receptor importin-α, which in turn recruits importin-β to mediate translocation through the nuclear pore. They find that mouse embryonic stem (ES) cells express the subtype importin-α1, whereas cells that have differentiated into neurons express importin-α5. Experimental manipulation confirmed that neural differentiation can be enhanced by combining the down-regulation of importin-α1 with the overexpression of importin-α5. Hence, the switching of importin-α subtype triggers neural differentiation of ES cells. The authors propose a mechanism by which importin-α subtypes function in either the undifferentiated or differentiated state by controlling the selective import of transcription factors into the nucleus—Oct3/4 for the former and Brn2 with SOX2 in the latter—which adds yet another layer of regulation for cell

CREDITS (TOP TO BOTTOM): SCHLIESSER *ET AL.*, *PHYS. REV. LETT.* **97**, 243905 (2006); DAVID STEPHENS/WWW.FORESTRYIMAGES.ORG

fate specification, this one acting via the intracellular trafficking of transcription factors. — BAP

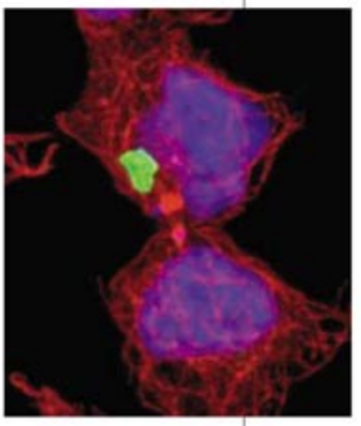
Nature Cell Biol. 10.1038/ncb1521 (2006).

CELL BIOLOGY

Sorting Out the Trash

When cells accumulate large quantities of proteins that have been damaged (for instance, via modification by reactive oxygen species) or that have not folded properly (for instance, as a result of mutations associated with neurodegenerative diseases), the degradative capacity of the intracellular quality-control system can be overwhelmed. Under these conditions, the aberrant proteins collect to form an aggresome, which is an inclusion body situated close to the microtubule-organizing center and just outside of the nucleus.

Rujano *et al.* examined the fate of cultured cells containing an aggresome, and of the aggresomes themselves, as the cells divided. Do aggresome-containing cells complete mitosis successfully? Are both daughter cells equally likely to inherit the parental garbage, or is one daughter preferentially spared? They found that aggresome-containing cells could indeed progress through mitosis productively and that the preexisting aggresome was inherited asymmetrically, yielding daughter cells relatively poor (or rich) in damaged proteins. Furthermore, a survey of cells in the epithelial crypts of the small intestine in two spinocerebellar ataxia (a neurodegenerative disorder) patients revealed a systematic allocation of the protein inclusions to the short-lived differentiated daughter cells, presumably ensuring the preservation of long-lived stem cells. — SMH



Aggregated protein (green) passes to only one daughter cell.

CREDIT: RUJANO *ET AL.*, *PLOS BIOL.* **4**, E417 (2006)

PLoS Biol. **4**, e417 (2006).

PSYCHOLOGY

Changing Attitudes

One emerging theoretical view posits two systems of reasoning: a slow-learning system that acquires and classifies associations over long periods of time, and a fast-learning module that emphasizes higher-order conscious cognition. A stimulus—for example, the negatively valenced word "hate"—can be paired in a subliminal fashion with a person's face (for example, Bob's); this association will induce subjects to regard Bob unfavorably, as assessed by their poststimulus choice of positive or negative adjectives, yet they will be unaware of having evolved this implicit attitude. Similarly, written descriptions of Bob's praiseworthy behavior will result in subjects expressing a liking for Bob, where this evaluation reflects a studied and thoughtful appraisal—that is, the formation of an explicit attitude. Rydell *et al.* show that these mental processes can be accessed separately and appear to operate independently. Not only are subjects capable of developing apparently inconsistent negative implicit attitudes and positive explicit attitudes about the same individual, but they can actually be influenced to invert their preferences by the subsequent presentation of subliminal (positive) words and supraliminal (negative) descriptions. — GJC

Psychol. Sci. **17**, 954 (2006).

CHEMISTRY

A Tale of Two Lattices

One promising aspect of metal organic framework (MOF) solids is the ease with which chiral components can be incorporated into the structural lattice. By linking metal centers with a network of chiral bridging ligands, researchers can prepare porous crystals with the potential to serve as robust asymmetric catalysts. However, assembling an extended MOF with specific steric and electronic properties remains highly challenging.

Wu and Lin highlight this challenge in presenting two MOF structures composed largely of the same building blocks, but exhibiting strikingly different lattice geometries and consequent properties. The first MOF was crystallized from a solution of cadmium nitrate and a chiral binaphthol derivative, appended with pyridines to bridge two metal centers. The hydroxyl groups in the lattice remained free to bind Ti(IV) centers, which in turn catalyzed ethylation of aromatic aldehydes with high yields and enantioselectivities. When a second MOF was prepared from the same precursors, but with the nitrate counterions replaced by perchlorate, a very different lattice structure emerged, which failed to catalyze the reaction. The authors suggest that steric crowding near the hydroxyls in this second structure inhibited effective binding of the titanium ions. — JSY

Angew. Chem. Int. Ed. **45**, 10.1002/anie.200602099 (2006).

www.sciencemag.org

Science

1200 New York Avenue, NW
Washington, DC 20005
Editorial: 202-326-6550, FAX 202-289-7562
News: 202-326-6500, FAX 202-371-9227

Bateman House, 82-88 Hills Road
Cambridge, UK CB2 1LQ
+44 (0) 1223 326500, FAX +44 (0) 1223 326501

Subscription Services For change of address, missing issues, new orders and renewals, and payment questions: 866-434-AAAS (2227) or 202-326-6417, FAX 202-842-1065. Mailing addresses: AAAS, P.O. Box 96178, Washington, DC 20090-6178 or AAAS Member Services, 1200 New York Avenue, NW, Washington, DC 20005

Institutional Site Licenses please call 202-326-6755 for any questions or information

Reprints: Author Inquiries 800-635-7181
Commercial Inquiries 803-359-4578
Corrections 202-326-6501

Permissions 202-326-7074, FAX 202-682-0816

Member Benefits Bookstore: AAAS/BarnesandNoble.com bookstore www.aaas.org/bn; Car purchase discount: Subaru VIP Program 202-326-6417; Credit Card: MBNA 800-847-7378; Car Rentals: Hertz 800-654-2200 CDP#343457, Dollar 800-800-4000 #AA1115; AAAS Travels: Betchart Expeditions 800-252-4910; Life Insurance: Seabury & Smith 800-424-9883; Other Benefits: AAAS Member Services 202-326-6417 or www.aaasmember.org.

science_editors@aaas.org	(for general editorial queries)
science_letters@aaas.org	(for queries about letters)
science_reviews@aas.org	(for returning manuscript reviews)
science_bookrevs@aaas.org	(for book review queries)

Published by the American Association for the Advancement of Science (AAAS), *Science* serves its readers as a forum for the presentation and discussion of important issues related to the advancement of science, including the presentation of minority or conflicting points of view, rather than by publishing only material on which a consensus has been reached. Accordingly, all articles published in *Science*—including editorials, news and comment, and book reviews—are signed and reflect the individual views of the authors and not official points of view adopted by the AAAS or the institutions with which the authors are affiliated.

AAAS was founded in 1848 and incorporated in 1874. Its mission is to advance science and innovation throughout the world for the benefit of all people. The goals of the association are to: foster communication among scientists, engineers and the public; enhance international cooperation in science and its applications; promote the responsible conduct and use of science and technology; foster education in science and technology for everyone; enhance the science and technology workforce and infrastructure; increase public understanding and appreciation of science and technology; and strengthen support for the science and technology enterprise.

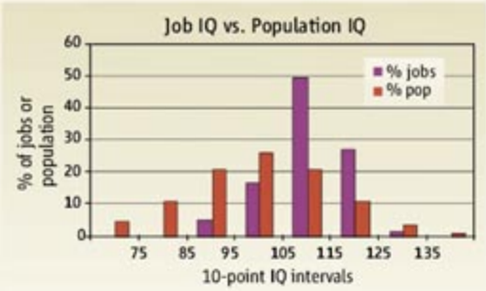
Information for Authors

See pages 120 and 121 of the 5 January 2007 issue or access www.sciencemag.org/feature/contribinfo/home.shtml

EDITED BY CONSTANCE HOLDEN

IQ Mismatch

This chart, says psychometrician Earl Hunt of the University of Washington, Seattle, shows why employers are so eager to automate. Many jobs requiring just-above-average cognitive abilities—say, 105 to 125 IQ points—are going begging, according to his recent analysis. In contrast, there is a surfeit of people with the potential to fill the relatively few jobs, such as Ph.D. physicist, that require supersmarts.



Hunt, who presented his data last week in San Francisco at the annual meeting of the International Society for Intelligence Research, based the chart on 706 jobs listed by the U.S. Department of Labor. Using detailed information available about IQ ranges of employees in 43 jobs, he extrapolated desirable levels of cognitive abilities for the rest and compared this with the IQ distribution of the general population. The dearth of skilled workers, notes Hunt, explains why, for example, car rental companies have equipped the person who checks in returning cars with a handheld computer that automatically calculates your fees. "The intellectual requirements of the job have been reduced," says Hunt.

Become an Evo Warrior

NET WATCH

What can you do if your local school board proposes a curriculum that downplays evolution? Or if your hometown newspaper runs an editorial supporting "intelligent design"?

This new site from the Federation of American Societies for Experimental Biology in Rockville, Maryland, offers advice and resources for scientists who want to defend Darwinism. Downloadable documents provide pointers on meeting with public officials, testifying at school board hearings, and related topics. Much of the advice is common sense, but some of it may be counterintuitive for scientists. For example, although you want your papers to run in prestigious journals, an op-ed will probably have more impact if it appears in the local paper than if it's accepted by *The Wall Street Journal*. The site also furnishes PowerPoint files on topics such as the importance of learning about evolution. >> www.evolution.faseb.org

CREDITS (TOP TO BOTTOM): SOURCE: EARL HUNT; NASA/JPL-CALTECH/UNIVERSITY OF ARIZONA; FASEB

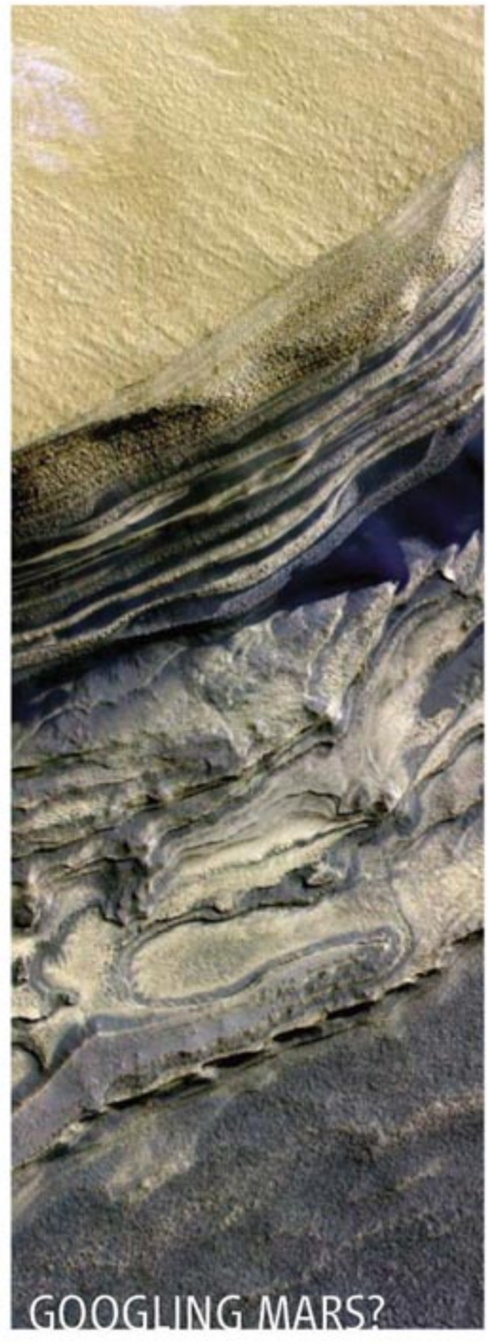
Cutting India's HIV Tally

A recent World Bank assessment of AIDS in Asia found India disproportionately afflicted, with 40% of Asia's population and 60% of its HIV infections. Official estimates of the number of Indians carrying the AIDS virus range from 5.2 million to 5.7 million.

But a new analysis suggests that these figures may exaggerate the problem by as much as one-third. Lalit Dandona of the Centre for Human Development at the Administrative Staff College of India in Hyderabad, reporting online 13 December in *BMC Medicine*, puts the number of HIV-positive Indians at closer to 3.5 million.

The National AIDS Control Organization (NACO) in New Delhi, the source of India's official estimate of 5.2 million infected, relies on a sampling method that involves monitoring AIDS clusters in certain large hospitals for a few months every year. Dandona and his colleagues took a different tack with a population-based random sampling method. They collected blood samples from 12,617 people in the southern Indian Guntur district and extrapolated that group's HIV-positive rate to Guntur's population of 4.5 million. The scientists came up with a total that was less than half of NACO's estimate for the region. Dandona's group also used Guntur's HIV prevalence to estimate the total number of infected in India's four southern states, which led to the new, lower-overall estimate for the nation.

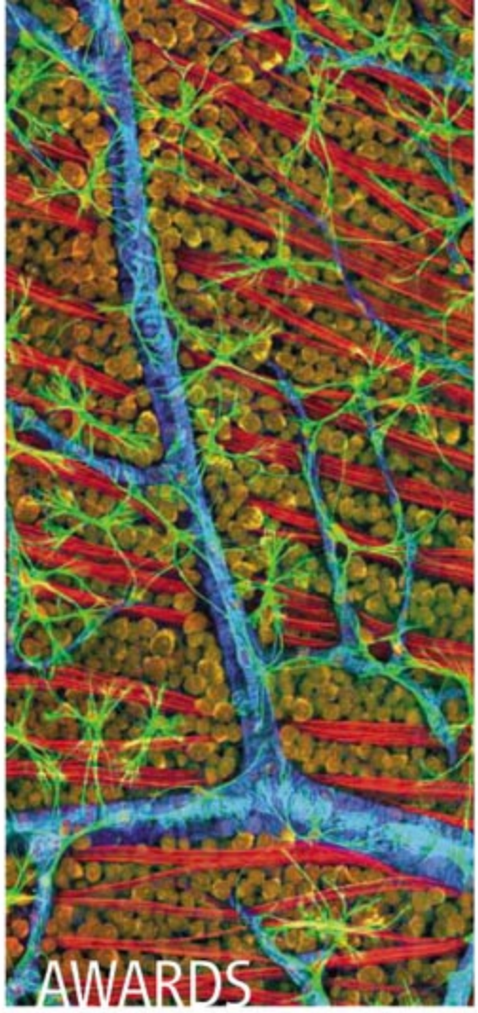
This is "very good news from a first-of-its-kind, robust study," says NACO Director General Sujata Rao, who says it shows that HIV infections in India are not spiraling out of control. The report is in line with a March 2006 paper in *The Lancet* by Rajesh Kumar and Prabhat Jha of the University of Toronto in Canada, who reported a one-third decline in new HIV infections in the worst-hit regions of India, thanks to condom use and AIDS awareness programs.



GOOGLING MARS?

Google recently inked a deal with NASA that will allow the Internet search engine firm to provide easy online access to images such as this new one from the latest camera to orbit Mars. The Mars Reconnaissance Orbiter spacecraft is slated to return more data than all previous Mars missions combined. This particular false-color image of a 700-meter-high, water-ice-rich cliff face in the martian north polar region reveals features, as small as a few meters across, that were shaped by past climate changes. The new finds include fine layering near the top of the cliff and house-size blocks of dirty ice emerging from lower layers.

EDITED BY KELLI WHITLOCK BURTON



AWARDS

IMAGE CONSCIOUS. Ansel Adams had Yosemite; Thomas Deerinck has the mouse retina. Deerinck, a staff scientist at the National Center for Microscopy and Imaging Research at the University of California, San Diego, has captured top honors at the Olympus BioScapes competition in San Diego for his image of the optic fiber layer in a mouse retina.

Deerinck is the first person to win both of the world's top microphotography prizes: In 2002, he took first place in Nikon's Small Worlds competition. Entries for BioScapes must depict the life sciences and be used in research, whereas Nikon's contest is open to all comers.

Deerinck's prize image, which earned him $5000 in Olympus products, is used in studies of neurofibromatosis, a disease that can cause blindness in children. It's on display at the San Diego Natural History Museum as part of a touring exhibit scheduled for Los Angeles, New York City, and other cities.

CREDITS (LEFT TO RIGHT): T. DEERINCK/THE NATIONAL CENTER FOR MICROSCOPY AND IMAGING RESEARCH/UCSD; GREG ADAMS, JEFFERSON LAB; COURTESY OF THOMAS FAMILY

Got a tip for this page? E-mail people@aaas.org

THE WORKFORCE

LIKABLE STRANGER. To not know the U.S. National Science Foundation (NSF) is to love it. That's the curious message from a new Gallup poll of 2600 U.S. adults asked about working for the federal government.

The Council for Excellence in Government (excelgov.org) is worried about what will happen when millions of baby boomers retire. So it asked Gallup to survey Generation Y (aged 18 to 29), older workers, and various white-collar professionals about the missions and attractiveness of 25 departments and agencies.

Overall, a bare 37% knew what NSF does, placing it ahead of only the near-invisible Office of Personnel Management. But the agency ranked fifth highest as a potentially interesting place to work. NASA scored near the top in both categories, second in interest and a lofty 86% in awareness. "That's a good place to start," says Gallup's Darby Miller Steiger. "But it means NSF needs to work harder on getting the word out."

PIONEERS

A LASTING GIFT. During the final weeks of their 9-year-old daughter's life, Shayne and Angela Thomas asked Children's Hospital of Philadelphia to develop a cell line from her drug-resistant neuroblastoma. Now, barely 3 months after Christi's death, scientists are gearing up for studies with the cell line, which could one day help others battle this childhood cancer.

The Thomases, of Tiffin, Ohio, received a crash course in drug development as a string of clinical trials kept Christi (top) alive for almost 4 years. Going the extra mile to create a cell line, her father says, "is the price I will pay" to help other families. So minutes after Christi died on 19 September, doctors drew a large volume of blood and shipped it to the Children's Hospital Los Angeles lab of C. Patrick Reynolds. Last month, the Thomases learned that the cell line, aptly named FU_NB06, is a reality. It should be available to scientists later this month.

AWARDS

LEIBNIZ PRIZE. Two women and eight men will receive $3.3 million each over 7 years as winners of this year's Gottfried Wilhelm Leibniz Prize for research. Presented by the German Research Foundation, the funds support work in diverse fields such as endocrinology and medieval history. American-born astrophysicist Guinevere Kauffmann will use the award in her work at the Max Planck Institute for Astrophysics on the Sloan Digital Sky Survey (SDSS), an ambitious project to create a 3D map of about 1 million galaxies and quasars. "The SDSS has been a tremendously successful and enjoyable project," she says.

Three Q's >>

In April, physicist **Fred Dylla**, 57, becomes executive director and CEO of the American Institute of Physics (AIP), which represents 10 professional societies and publishes a variety of journals. An administrator at the Thomas Jefferson National Accelerator Facility in Newport News, Virginia, Dylla will succeed the retiring Marc Brodsky.

Q: What is the biggest challenge facing AIP?
Our primary challenge is to fully embrace and push for the recommendations in *Rising Above the Gathering Storm*, the [National Academies] report that calls for increased funding for the sciences and science education.

Q: Will the AIP journals move toward open access?
Of course, we want the journals to be widely accessible. But the community also wants any publication to be high-quality, peer-reviewed, and archival, and those things have to be paid for. I think there is a business model emerging in which publication fees from the author and subscription fees from large institutions will pay for the value added.

Q: What can AIP do to increase diversity in physics?
There's no silver bullet. You have to address the entire pipeline from grade school to mentoring professionals.

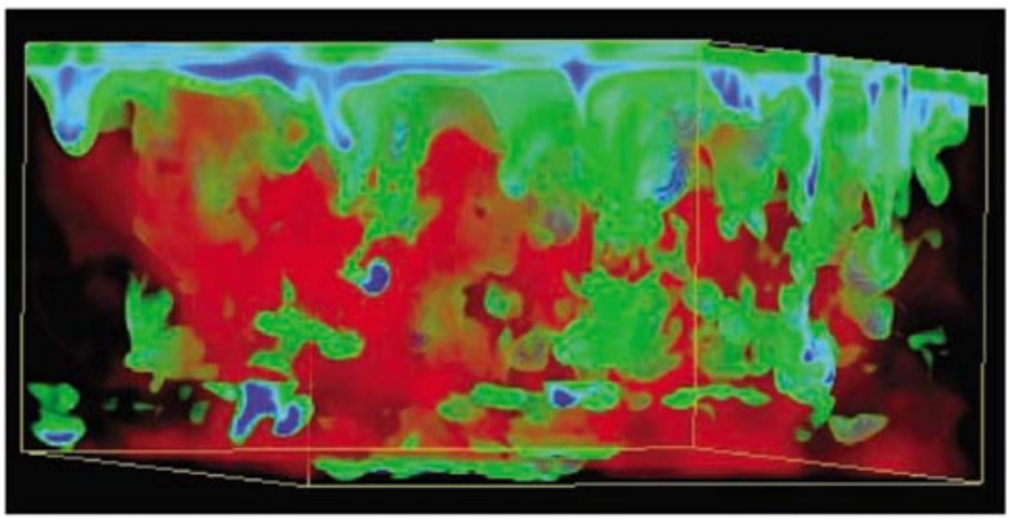
Focus on autism 27

The House's "Mr. Science" 28

2007 U.S. BUDGET

NSF Braces for Opportunities Lost

Wayne Pfeiffer and his colleagues at the San Diego Supercomputer Center didn't mind working over the holidays on a proposal due 2 February to the National Science Foundation (NSF). They knew their counterparts at other NSF-funded supercomputing centers would be doing the same thing. And besides, the prize seemed worth the extra effort—a $200 million machine capable of performing at the petascale level (10^{15} operations a second) and, with it, leadership of the next generation in supercomputing.



Hot modeling. This supercomputer simulation helps scientists understand convection and magnetic flux at the sun's surface. A petascale supercomputing initiative may be caught in the budget freeze.

What they didn't figure on, however, is that there might be no winner at all. That outcome has suddenly become quite probable, at least for 2007, after the outgoing Republican Congress adjourned without finishing work on the federal budget and the incoming Democratic leadership announced its intention to freeze spending at current levels (*Science*, 15 December, p. 1666) until October. A delay—or worse—in the petascale computing competition is only one of dozens of unhappy scientific consequences of the current legislative train wreck for NSF, which had high hopes for an 8% boost this year in its $5.6 billion budget. As the 110th Congress convenes this week, some science lobbyists are hoping against hope to salvage a piece of what was supposed to have been a banner year for the agency.

NSF was one of three entities—along with the Department of Energy's Office of Science and the in-house labs of the National Institute of Standards and Technology—targeted for significant increases as part of the Bush Administration's American Competitiveness Initiative (ACI). The president's 2007 budget request to Congress last February contained the first installment of what was intended to be a 10-year doubling of federal basic research spending in the physical sciences.

NSF, which currently is able to fund barely one in five proposals, is already inundated with good suggestions about how to spend its money. Its proposed 2007 budget was chock full of fresh ideas, from a new $98 million Arctic research vessel to a $25 million pot of money to fund "frontier" research at the intersection of engineering and a host of other disciplines. Some would involve global activities, such as NSF's planned $61 million boost for polar research to take advantage of the International Polar Year (IPY) that begins in March and runs until early 2009. And all were predicated on robust annual growth in NSF's budget.

The winner of the petascale computer competition, for example, was scheduled to receive a $50 million downpayment in 2007, with similar increments in each of the next 3 years. Any slip in the 2007 start date, say computer scientists, could put a crimp in the entire project. "We had developed a pretty detailed timeline with milestones" for petascale applications, pilot studies on how to run them, and potential collaborations, explains San Diego's Allan Snavely, director of the center's performance modeling and characterization lab. "Now that momentum could stall. In addition, collaborative teams can't sit around and twiddle their thumbs for a year. Some will move on other things."

NSF officials say they are sticking to the original schedule for the competition, including site visits this spring. But if the agency's budget remains flat, they will have to make an award sometime after the 1 October start of the 2008 fiscal year. And the project's long budget tail will require sustained increases through 2011, warns Thom Dunning, director of the National Center for Supercomputing Applications in Urbana, Illinois, which is also competing for the petascale machine.

The competition within the engineering directorate for "frontier" research had already attracted 257 preproposals, and NSF was planning to invite 50 teams to submit full proposals by spring and to make 11 awards. The money represents more than half of the directorate's expected 6% increase for 2007. With a flat budget, however, the new office will be lucky to fund more than a couple of projects.

With regard to the Arctic vessel, time is money. A solicitation went out in October to build the ship, and any delay in construction will add to the final cost. Geoscientists say it will likewise slow their quest to understand the impact of global warming on this bellwether region of the planet.

Social scientists also hoped to make a splash in 2007 by ramping up a program to develop what NSF calls the science of science and innovation policy. In response to a complaint by presidential science adviser John Marburger that too little is known about what drives innovation and how to measure it, NSF's social, behavioral, and economic sciences directorate was ready to plow half of its proposed $14 million increase into three research competitions. Anticipating a flat budget, directorate head David Lightfoot ▶

CREDIT: ROBERT STEIN AND AAKE NORDLUND/NCSA/NSF

says that "if we decide to put out a solicitation, it will be a lot smaller."

With so much on the line, some science lobbyists are hoping that legislators will allow some melting around the edges in the year-long budget freeze they are expected to pass before the current spending resolution expires on 15 February. But most are setting their sights on the 2008 budget cycle, which begins at the end of the month with the president's State of the Union address and, a few days later, his budget request to Congress. Toward that end, three Senate advocates of legislation to authorize ACI spending levels have asked President Bush "to continue to make this issue a top priority in your budget and for your administration." The 21 December letter from Senators Pete Domenici and Lamar Alexander, Republicans from New Mexico and Tennessee, respectively, and Senator Jeff Bingaman, a Democrat also from New Mexico, could well have been written by the scientific community itself: "If America is going to continue to be the global economic leader, we can not afford to let this wait." **–JEFFREY MERVIS**

ENDANGERED SPECIES

U.S. Weighs Protection for Polar Bears

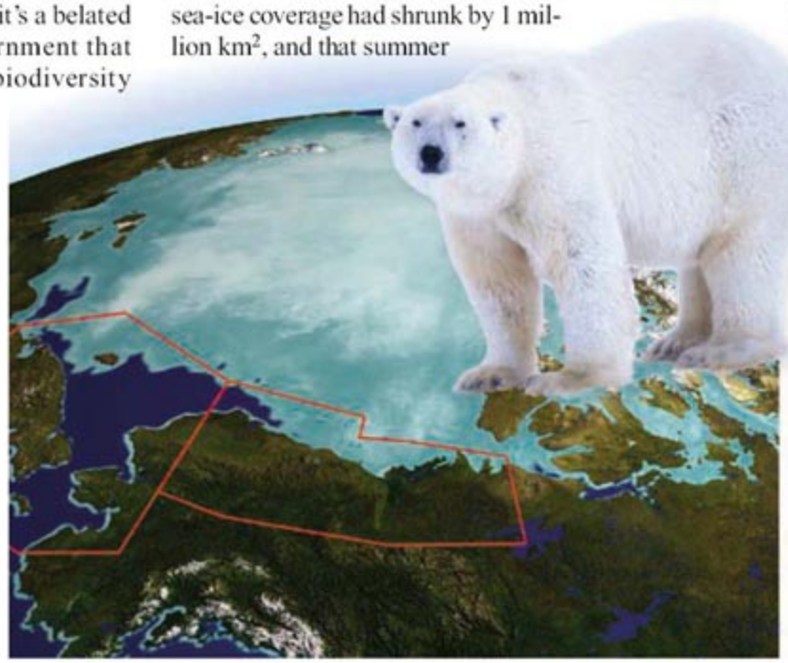
Weighing up to 800 kilograms, the polar bear is king of the Arctic. And last week, the U.S. government made it the poster child for global warming.

Citing warming temperatures that are melting the sea ice that polar bears call home, the U.S. Fish and Wildlife Service announced that within the next year it will decide whether to protect these animals under the Endangered Species Act. The move represents the latest in a multidecade international effort to maintain the species. And for environmental groups, it's a belated recognition by the U.S. government that global warming is affecting biodiversity and ecosystem health. "It's a really watershed moment," says Andrew Wetzler, an attorney at the National Resources Defense Council (NRDC) in Washington, D.C., which has been pushing for polar bear preservation.

The world's 20,000 to 25,000 polar bears are divided into 19 populations distributed across the Arctic. Throughout the winter, they hunt seals from sea ice that expands southward each winter and contracts as the temperature rises. Often stranded on land in the summer, the animals fast.

Arctic nations have worked for decades to control the hunting of polar bears. But the dangers to the bears' habitat are a more recent concern. In 2005, the Center for Biological Diversity, an environmental advocacy group based in California, asked the Fish and Wildlife Service to add polar bears to the list of threatened species because of fear that continued habitat loss might drive them into extinction. Last week's announcement came after the center, Greenpeace, and NRDC took the agency to court for not responding to the request.

There is growing evidence that polar bear habitat—sea ice—is declining, with less of it forming every year. In 2004, the eight-nation Arctic Climate Impact Assessment concluded that the Arctic was warming twice as fast as the rest of the world, that the average annual sea-ice coverage had shrunk by 1 million km^2, and that summer sea ice could disappear by 2100 (*Science*, 5 November 2004, p. 955). In September 2005, Julienne Stroeve of the National Snow and Ice Data Center in Boulder, Colorado, reported that satellite data showed sea-ice coverage had reached an all-time low and was shrinking at an annual rate of about 8%.



Slipping away. As the extent of sea ice (white) declines, the two populations of U.S.–based polar bears (red lines) will have a harder time surviving.

CREDITS: NASA; PHOTOS.COM

A 2004 study of polar bears in Canada's Western Hudson Bay—where sea ice is breaking up 3 weeks earlier than it had 30 years ago—highlights the problem for the bears. It found fewer than 1000 bears, down from 1200 a decade ago. The bears are also thinner, and fewer cubs are surviving. "As sea ice melts, bears are having a much more difficult time," says Andrew Derocher, a population ecologist at the University of Alberta, Edmonton. Researchers don't have firm numbers on the population of bears in the southern Beaufort Sea, which extends across Canada and the United States, but the bears there are also thinner and have fewer cubs, suggesting they, too, are in trouble. In June 2005, these findings prompted the World Conservation Union (IUCN) to list polar bears as "vulnerable" on its Red List of endangered species.

The U.S. announcement provides for a 3-month comment period. Only if the agency decides to list the polar bear as threatened will it examine how best to care for the species. Wetzler says that's a necessary first step to controlling the greenhouse gas emissions that are the real cause of the bears' plight. "The polar bears' survival is going to depend on our ability to come to grips with global warming," he says. **–ELIZABETH PENNISI**

SCIENTIFIC MISCONDUCT

Japan's Universities Take Action

TOKYO—A series of outstanding scientific misconduct cases ended suddenly and decisively in recent days: Two leading Japanese universities fired scientists because of questionable publications and a researcher is reported to have resigned from a third university over alleged mishandling of research funds. Although university officials say the timing was coincidental, researchers say the unprecedented actions suggest that Japanese institutions are now taking a tougher line on scientific integrity.

The resolution of the cases "takes a thorn out of our hearts," says Norihiro Okada, a molecular biologist at the Tokyo Institute of Technology who with other scientists challenged work by one of the three groups. He says that bad research tarnishes the image of the entire scientific community, adding that "the recent actions of these universities are very much welcome."

Abrupt exit. Kazuko Matsumoto (*left*) was accused of mishandling funds; Kazunari Taira was dismissed for "negligent research conduct."

On 20 December, Osaka University announced that it had fired a professor previously found by an investigating committee to have fabricated research data. In a brief statement posted on the university's Web page, President Hideo Miyahara confirmed the decision but did not identify the professor or give any details. However, Japanese newspapers identified the scientist as chemist Akio Sugino and the questionable paper as having appeared in the *Journal of Biological Chemistry*. The journal's Web site lists a paper co-authored by Sugino and published in July as having been withdrawn in August by the authors. According to a local press report, one or more co-authors contacted a departmental research fairness committee, which concluded that Sugino acted alone in fabricating data for at least one paper, although critics also raised questions about other papers. One of the co-authors committed suicide last September, but university authorities declined at the time to comment on a possible connection to the research misconduct allegations.

Waseda University in Tokyo concluded an investigation into an alleged misuse of research funds by posting a report on its Web site on 19 December. The report noted that a professor had allegedly drawn money from public research grants to pay students for part-time work but diverted the money into a private bank account. Someone within the school tipped off university officials earlier this year. The professor was previously identified as Kazuko Matsumoto, who has maintained the funds were used for legitimate research purposes (*Science*, 7 July, p. 31). In its report, the university said an investigating committee found no evidence that the professor (whom it didn't name) had embezzled research funds but noted that the professor was "aware" that payment requests had been mishandled. Two university officials, including the president, received salary cuts of unspecified amounts, and two other officials were given official warnings for lax oversight. In addition, the university will return roughly $1.8 million in research funds to the government. A local newspaper separately reported that Matsumoto resigned on 22 December. She could not be reached for comment.

Finally, on 27 December, the University of Tokyo dismissed Kazunari Taira and Hiroaki Kawasaki for "unreliable" research practices involving papers that appeared in *Nature*, *Nature Biotechnology*, the *Proceedings of the National Academy of Sciences*, and other journals. The RNA Society of Japan raised doubts about the work in an April 2005 letter to the university requesting an investigation. A university investigating committee found Taira's group did not have raw data or notebooks to support results for a number of experiments focusing on RNA. Taira, a chemistry professor, and Kawasaki, a research associate, claimed raw data entered directly into a computer had been lost. The committee later concluded there was no evidence the experiments could be reproduced. (*Science*, 23 September 2005, p. 1973; 3 February, p. 595; 17 February, p. 931) In a statement posted on its Web page on 27 December, the university said it could not prove deliberate fraud but was dismissing the pair for discrediting the university through "negligent research conduct."

For both Tokyo and Osaka universities, it was the first time faculty members had ever been dismissed for alleged scientific misconduct. Scientists say that previously, researchers under a cloud of suspicion would have been quietly asked to seek another job. But as funding and competition for funding have grown, scientists—and taxpayers—are demanding greater transparency and accountability.

Over the past year, all three of these universities introduced codes of conduct for researchers and established offices or committees to promote good ethics and investigate allegations of fraud.

But the wider scientific community may not recognize the need for enforcement. A survey conducted over the past year by the Science Council of Japan, the country's largest association of researchers, found that just 13.3% of 1323 responding institutions had adopted a code of ethics and only 12.5% had established procedures for handling allegations of misconduct.

Makoto Asashima, a developmental biologist at the University of Tokyo who chaired the council's Committee on the Code of Conduct for Scientists, says the recent announcements by the three universities will likely spur other institutions to take action. His committee drafted a suggested code of conduct and recommendations for countering misconduct. "Each university and research institution should draw up and implement its own procedures and policies regarding scientific conduct," he says. This would help establish public trust in research institutions.

–DENNIS NORMILE

CREDITS (LEFT TO RIGHT): IUPAC; D. NORMILE/*SCIENCE*

BIOMEDICAL POLICY

New Autism Law Focuses on Patients, Environment

Congress has told the National Institutes of Health (NIH) to pick up the pace of its research on autism, with an emphasis on early diagnosis, treatment, and the role of environmental factors. The Combating Autism Act, passed in the waning hours of the 109th Congress and signed into law 19 December by President George W. Bush, authorizes a major increase in spending and orders NIH to come up with a detailed research plan for making progress in understanding and treating the disorder.

"It's giving us a flashing green light to move faster on autism," says Tom Insel, director of NIH's National Institute of Mental Health in Bethesda, Maryland. What reauthorization bills don't provide, however, is any money. And with most government agencies preparing for flat budgets in 2007 (see p. 24), Jon Retzlaff of the Federation of American Societies for Experimental Biology, says it's "inconceivable" that legislators will divert scarce NIH dollars to autism.

NIH estimates that it spent $101 million last year on autism-related research. The new law allows that figure to increase to $132 million this year and to $210 million by 2011. In addition, the Centers for Disease Control and Prevention in Atlanta, Georgia, which focuses on the epidemiology of autism, could grow its programs from the current budget of $15 million to $21 million by 2011.

Advocates for autism research hope that things will speed up even without an immediate funding boost. For example, an interagency committee that coordinates autism research must now submit for the first time an annual report on progress in causes, diagnosis, and treatment. "That certainly puts more emphasis on its role," says Manny DiCicco-Bloom of the Robert Wood Johnson Medical School in Piscataway, New Jersey, and a member of the board of directors of the nonprofit Autism Speaks. "Perhaps with more teeth, [the committee] can make real changes in policy and levels of performance."

The new law also orders the committee, which reports to the Health and Human Services secretary, to create and implement a strategic plan for autism research. Advocates say that's much better than a 2003 report from the committee, a list of short- and long-term goals with no rankings or recommendations on how to carry them out. The plan must be updated every year and include a draft budget for accomplishing research goals. Insel, who chairs the committee, says he's already convened a working group. He hopes the plan will be ready by the summer.

Although the law doesn't set any specific funding levels, it directs NIH to expand, if funds are forthcoming, its work on diagnosis, treatment, and possible environmental causes of autism. That's music to the ears of Jon Shestack of the advocacy group Cure Autism Now, who says that NIH hasn't done nearly enough on this front.

In November, an ad-hoc NIH review committee agreed and recommended investigating the possible role of neurotoxic compounds such as pesticides and mercury, developing new biomarkers for exposure, and studying exposure in pregnant women with autistic children. Insel and other scientists agree that those topics are important but argue that, absent more money, NIH should stick with its existing programs. "The research agenda is excellent," says epidemiologist Eric Fombonne of McGill University in Montreal, Canada.

–ERIK STOKSTAD

New focus. President George W. Bush signs into law new marching orders for understanding and combating autism.

CREDIT: ERIC DRAPER/WHITE HOUSE/HANDOUT/CNP/CORBIS

SCIENCESCOPE

Japanese Budget Sags

Japan's spending on research is poised to drop for the third year in a row. The debt-plagued government has budgeted $29.5 billion, down 1.8%, for science spending in the fiscal year beginning 1 April. Losers include RIKEN's Radioactive Isotope Beam Factory, whose $23 million amounts to a 28% dip and about half of what had been requested for the just-completed exotic isotope accelerator. Parliament is expected to make minor changes before signing off soon.

Not all the research news is grim for scientists. Funding for competitively reviewed grants will grow 1.4% to $4 billion. In addition, a supplemental budget provides money primarily to strengthen the earthquake resistance of scientific facilities. Kiyoshi Kurokawa, science adviser to the prime minister, notes that combining the two budgets results in a net 2.3% rise in science-related spending over last year. But critics contend the supplemental budget simply steers money into the politically powerful construction sector.

–DENNIS NORMILE

Rovers Reloaded

A New Year's resolution shared by NASA's Spirit and Opportunity: Think more for myself. The pair of weary Mars explorers have received a software upgrade to allow them to recognize dust devils and clouds and select only relevant sections of the images to transmit to Earth, freeing up communication time and manual labor for scientists. Other new features include better obstacle avoidance software. The rovers' missions are entering their fourth year. –ELI KINTISCH

Exhibiting Restraint

Plans to build a new government-funded science and technology museum in Ottawa have been undercut by Canada's top treasury official. The Canada Science and Technology collection is currently dispersed among three buildings, including a former bakery. The long hoped-for building would bring the 36,000-item collection, which includes Canadian-made satellites, antique scientific and medical instruments, and nanotechnology exhibits, under one roof along with curators, researchers, and cataloguers. Museum officials cut back their proposal last year from $600 million to $400 million. But even the smaller figure is too much for Treasury Board of Canada president John Baird, who cited more important taxpayer needs in an interview with the *Ottawa Citizen* newspaper last week. A newly appointed museum chair will now reevaluate the situation. –PAUL WEBSTER

Wedded to Congress. Bart Gordon delayed marriage and family until his 50s to pursue a career in public service.

CREDIT: TOM WILLIAMS

PROFILE: BART GORDON

New Chair of House Science Panel Takes Extreme Route to Moderation

Representative Bart Gordon (D–TN) thinks of himself as a moderate Democrat. "On fiscal matters, I'm conservative, and on personal liberties, I'm more liberal," says the incoming chairman of the House Science Committee, which he is already touting as a user-friendly panel "of good ideas and consensus."

But the word moderate hardly describes his fiercely competitive nature, or how the 59-year-old lawyer, born and raised in the Middle Tennessee district that he has represented for 22 years, lives and breathes politics. Those traits, along with his close ties to Representative Nancy Pelosi (D–CA), the new House Speaker, and her promise to make innovation a centerpiece of the Democratic agenda, could elevate the status of the traditionally low-profile panel and make Gordon a significant player in the 110th Congress that opened for business this week.

If that happens, it won't catch Gordon by surprise. He first sketched out his political future as a high school senior in Murfreesboro, Tennessee, while working on the political campaign of a family friend, and the blueprint hasn't changed in 40 years. "I decided then and there to go into public service. And as the son of a farmer and schoolteacher, I felt that Congress was probably the highest office I could achieve with just hard work and some degree of certainty. And so I spent the next 18 years preparing to do that." His 80-year-old mother, Margaret Gordon, recalls that her son "didn't have time for hobbies" as a child. "He's so focused it's pathetic," she jokes.

How focused is Bart Gordon? He admits that he didn't marry until his 50s because of the demands of his job. When Gordon decided to compete in an annual 5k charity race that pits politicians against the Washington media who cover them, he asked the nationally recognized track coach at his alma mater, Middle Tennessee State University, for advice on his workouts. It worked: Gordon holds the unofficial title of "fastest member of Congress."

Going at less than full speed just isn't in his nature. Asked whether he ever thought of taking a more relaxed approach to life, and to his service in Congress, Gordon shakes his head. "This is a fast track. … I would want to excel in whatever I do. It's just not any fun to be in the middle of the pack."

Thanks to Democratic electoral victories in November, Gordon will have the chance to lead a panel that oversees the lion's share of the government's nonmedical civilian research activities. His to-do list comes from the mainstream of his party—strengthen U.S. competitiveness, develop greener sources of energy, improve science and math education, and keep a close eye on the Bush Administration's management of the federal research enterprise. But his legislative strategy is so straightforward that it comes across as radical. "To me, a good idea is a good idea," he says. "Rather than taking 5 or 6 years to put together a massive piece of legislation like a telecommunications or an energy bill, I think we should try to develop a consensus on the good idea and move ahead with it."

For Gordon, moving ahead on a good idea meant making a bid for Al Gore's House seat in 1984 when Gore decided to run for the Senate. Since then, Gordon has been reelected 11 times, usually by comfortable margins, despite an increasingly suburban district that tends to vote Republican. "He's pretty well convinced his potential opponents, and the Republicans, that they should do their mining somewhere else," says longtime friend and political confidant Andy Womack, a State Farm insurance agent in Murfreesboro.

Although Gordon says he has no ambitions for higher (read governor or senator) office, that doesn't mean he lacks a global vision. However, ask him whether the United States can hold its own against the growing technological prowess of China and India, and his answer couldn't be closer to home. "I've got a 5-year-old daughter who I really believe could be part of the first generation of Americans who could inherit a standard of living lower than their parents," he says, as his throat catches and a tear forms in the corner of his eye. "That's a complete reversal of the American dream."

In the midst of moving both his personal and science committee staffs on Capitol Hill last month, Gordon spoke with *Science* about his political philosophy, science, and his plans for the committee. Here are excerpts from that interview.

On passing an innovation bill

"I realized that we have some jurisdictional problems over here in the House. So what I told Senator [Lamar] Alexander (R–TN) is that they should get their bill out as quickly as possible. And then rather than have a parallel bill, we'll come out with a bill that falls within our [narrower] jurisdiction. Then in conference we can put the two bills together. Yes, Humpty Dumpty can be put back together again. I'd rather do that than to slow this thing down by 2 or 3 years by trying to pass exactly parallel bills in each house."

On working with appropriators

"I don't think there has been adequate communication between authorizers and appropriators. After 22 years, I think I

have pretty good relations with both Democrats and Republicans. The appropriators have the dilemma of unlimited wants and limited amounts of money. But I think we can sit down and talk about priorities. In fact, I think it would be interesting to have some joint hearings. It needs to be a collaborative effort. Now, that doesn't mean you get everything you want. But it does mean that you agree to make the best out of limited resources.

"You have to do more than just say, 'We need more money,' or that 'the National Science Foundation needs to be doubled'—I'd like 5 years, but 7 years is probably more realistic. We have to sit down and do some give and take. Within the NASA budget, I suspect that whatever we do, there won't be adequate funds to do everything that NASA has been charged with doing."

On legislative oversight

"I think that the science committee, and Congress as a whole, has acquiesced in its oversight responsibilities. And I think that if somebody is not looking over your shoulder, you become cavalier. I saw it happen to the Democrats [when they controlled Congress prior to 1995]. If you recall, the science committee, under the Republicans, did away with the oversight committee, which was our only vehicle for those investigations. At the same time, I sincerely think that the Republicans were stifling some scientific conclusions and looking to staff committees with people who would go along with those conclusions and discourage any opposition.

"I think accountability is important, to save taxpayer dollars and get the most out of government programs. I also think, quite frankly, that we need to do a better job of reviewing whether or not the Administration is cooking the books with science, and prejudging its findings. I think that will be less likely to occur if somebody is looking over their shoulder. My purpose is not to embarrass someone about their prior activity, but rather to it make it clear that from now on we will be providing oversight, so don't do it anymore."

> "I think that we need to do a better job of reviewing **whether or not the Administration is cooking the books with science.**"
>
> —Bart Gordon

On a new agency for energy research

"I think we should follow the DARPA [Defense Advanced Research Projects Agency] model, so there would be fewer strings. Plus, the idea of an ARPA-E [Advanced Research Projects Agency, Energy] would be to have a better focus. You can't do a hundred things. We're hoping to find the best seven or eight approaches to renewable energy, and then focus on them. Bring together the national labs, the public and private sectors, to focus on the problem. And I think that the Department of Energy needs a little encouragement to get that done. ... I've talked with [Energy] Secretary [Samuel] Bodman. And his reaction is status quo. But status quo isn't getting the job done."

On sources of advice

"The science committee has a long-tenured, well-credentialed staff, and I feel very comfortable with their advice. … I'm also frequently meeting with associations and university presidents. Of course, as my grandfather used to say, 'The most important road in the county is the one in front of your house.' So you need to apply a little bit of a filter to what they say. But I find that in the scientific community, there really aren't monetary drivers as much as passions that people have."

On global competition

"There are seven billion people in the world, and half of them make less than $2 a day. We can't compete against $2-a-day labor, and we wouldn't want to. But now India and China and other countries are also investing in R&D and starting to combine their cheap labor with innovation. So in order to maintain our standard of living, we have to increase our productivity even more.

"We want to develop the technology to be first to market, time and again. But we also need a workforce that can work at a high skill level, and not just based on recruiting the one in a hundred students who wants to be a scientist."

On the Administration's scientific team

"I think NASA Administrator [Michael] Griffin is certainly one [of the most impressive]. Partly because who he followed [Sean O'Keefe], and partly because he is both knowledgeable and candid. We don't always agree, but you know you'll get a honest, from-the-gut assessment. We don't always get that.

"Having said that, however, I think that this is a top-down Administration, and there's a lot of pressure from the top down to make the conclusions match the preconceived notions of the Administration. … I think that [presidential science adviser John] Marburger would say that he hasn't been constrained. But I think we need to look into that more. I think he's an honorable and capable man. But he's under a lot of pressure, too."

–JEFFREY MERVIS

CREDITS: TOM WILLIAMS

CREDIT: GETTY IMAGES

Indonesia Taps Village Wisdom to Fight Bird Flu

Participatory epidemiology is Indonesia's first step on a long road to controlling avian influenza

SUKASARI, INDONESIA—Sobandi and Rahmet Hidayat stroll down a dusty lane in this village in the highlands of central Java, stopping to chat with a group of women escaping the midday sun on the porch of a modest stucco house. After doffing their shoes and offering the traditional two-handed greeting all around, the two animal health officers sit on the tiled porch floor and ask about chickens. As children gawk at the visitors and a lone hen scratches in the front yard dirt, the women describe a big die-off that occurred last January. The vets already knew about that one; it had wiped out virtually all poultry in the entire district.

It's the dry season, so instead of cultivating rice, most men are off working as day laborers or tending roadside produce stalls. Soon an elderly man joins the conversation and mentions a second, smaller die-off this past July. This is new information. The villagers hadn't thought to report it and weren't sure to whom to report such problems. As the chicken struts across the porch and the children return to their play, Sobandi and Rahmet Hidayat, who like many Indonesians use only given names, ask about symptoms to see whether the villagers can distinguish the highly pathogenic H5N1 avian influenza from other diseases. The old man gives what could be a textbook description of the symptoms—swollen and bluish discoloration of combs and wattles, lethargy, and then sudden death—that characterize H5N1.

After another round of handshakes, the animal health officers continue down the lane to a second group of women peeling cassava on a porch shaded by banana and coconut trees. As the call to noon prayers comes from the village mosque, the vets squat on their haunches and ask the women about their chickens.

This scene has been repeated innumerable times across Java over the past year as Indonesia has struggled to gauge the extent of its avian influenza problem, by all accounts the worst in the world. The virus is endemic among poultry throughout much of the country. "It is very serious; 30 of 33 provinces in Indonesia are infected by avian influenza," says Bayu Krisnamurthi, who heads the Indonesia National Committee for Avian Influenza Control and Pandemic Influenza Preparedness. "You simply couldn't get more virus in the environment," says Jeffrey Mariner, a veterinarian at the International Livestock Research Institute in Nairobi, who is helping train surveillance teams like this one under the auspices of the Food and Agriculture Organization (FAO) of the United Nations.

The close interaction of infected poultry and humans inevitably leads to human cases, and Indonesia has the worrisome distinction of having the highest number of H5N1 human fatalities of any country, curiously clustered among blood relatives (see sidebar on p. 32). Experts fear that Indonesia provides the perfect setting for H5N1 to evolve into a form easily passed among humans, touching off a global pandemic. Says World

Health Organization virologist Keiji Fukuda: "Reducing infections in poultry is a critical aspect of reducing the risk to people." But, as yet, adds Peter Roeder, an FAO animal health officer, "there is still no systematic control program."

The first step in such a program is to track where and when the outbreaks are occurring, especially among the chickens kept in backyards by 60% of all Indonesian households—an estimated 300 million birds. That's the challenge for a new approach, called participatory epidemiology, that Sobandi and Rahmet Hidayat are trying. Pioneered by Mariner, formerly a professor at Tufts University's International Veterinary Medicine Program in North Grafton, Massachusetts, and Christine Jost, a Tufts assistant professor, it sounds simple enough: Train teams of vets to tap into local knowledge of where and when outbreaks are occurring, and then enlist villagers' cooperation in control efforts. The basic fieldwork provides epidemiological data on how the disease is spreading and kept in circulation, which in turn leads to higher-level strategies for control.

But participatory epidemiology has never been tried for avian influenza before—and never on this scale for any disease. Even so, international and Indonesian animal health officials believe it will be a key component of bringing the H5N1 crisis under control, both here and throughout the developing world.

Starting from scratch

"When we started, we had no idea where the disease was and how much of it there was," says FAO's Roeder. One reason is that Indonesia's animal health infrastructure had deteriorated badly over the past decade, in part, a victim of its own success. John Weaver, a senior FAO adviser in Jakarta, says that after Indonesia brought scourges such as foot-and-mouth disease under control in the early 1990s, vet services became "an easy target" for cost-cutting during the Asian financial crisis at the end of the decade. To reduce national expenditures, animal health services were turned over to local control. The fragmentation and disruption hit the country's poorer regions particularly hard. As an example, Asjachrena Lubis, a Ministry of Agriculture official, points to Maluku Province, a group of islands in eastern Indonesia. It had five government vets under the national system. But with localization, they all left for better paying jobs on Java. And now, "as far as we know, Maluku Province doesn't have a single [government] vet," she says.

Even where local governments have managed to continue animal health services, limited staffing means vets only respond to reports of problems. And even that support never extended to backyard flocks. "For those who keep a few chickens, it's not worth it to call someone who is a 2-hour motorcycle ride away. And what will they do when they get there?" asks Alison Turnbull, a former Tufts student who is helping train the new surveillance teams.

Participatory duo. Veterinarians Christine Jost and Jeffrey Mariner pioneered participatory epidemiology to counter rinderpest in Africa and are now applying the technique to bird flu in Indonesia.

Mariner had faced a similarly underdeveloped vet infrastructure when he was a Tufts veterinary graduate student working for FAO's Global Rinderpest Eradication Program in Africa in the early 1990s. "I realized that the farmers knew a lot more about where rinderpest was than the veterinary officials," Mariner recalls. He adapted older community-based schemes to have vets ask herdsmen about animal deaths and illnesses. This participatory surveillance enabled authorities in Sudan to target vaccination programs that eradicated rinderpest from the country. The technique was later used for rinderpest in Pakistan and for hog cholera in South America.

As secretary of FAO's rinderpest program, Roeder was intimately familiar with participatory surveillance and suggested it when he came to Jakarta in the summer of 2005 to help Indonesia craft a response to bird flu. Indonesian health officials agreed it would be a good fit. Krisnamurthi, of the national avian influenza committee, says, "Indonesia is not just vast geographically but also socially and culturally." The country's 220 million people are spread over 3000 inhabited islands and represent some 350 ethnic populations. Many of these groups are wary of government programs, he says. Roeder adds that Indonesia lacks the strong central government and established veterinary capabilities that enabled top-down bird flu control programs to work in Thailand, which relied on aggressive culling, and Vietnam, which introduced massive vaccination.

FAO once again turned to Mariner and Jost to help set up the program. Early in 2006, with $1.5 million in funding from the U.S. Agency for International Development, they established a pilot program in 12 districts in Java that still had some publicly funded vets, forming teams of two specializing in either participatory disease surveillance or participatory disease response.

From early on, it was clear that villagers were well aware they were dealing with an unusual—and unusually lethal—disease, says Jost. They were calling it "new" Newcastle or "strong" Newcastle disease or just "plok"—the sound of a dead chicken falling from a perch. And each week, the teams began reporting one or two previously undetected outbreaks. "It turned up much more avian influenza than anyone expected," Mariner says. "Poultry populations were fully saturated."

Can we talk? To track outbreaks of the highly pathogenic H5N1 avian influenza, Indonesia's vets are going door to door asking villagers about outbreaks among their poultry.

CREDITS (TOP TO BOTTOM): D. NORMILE/SCIENCE

Not just chickens. The prevalence of H5N1 among poultry throughout Indonesia has inevitably resulted in a rising number of human cases and fatalities.

Human Cases Create Challenges and Puzzles

One day in mid-September, a 23-year-old Indonesian man bought four steeply discounted dead chickens at a poultry market in the central Java town of Bandung, carried them home in a plastic bag, and together with his 20-year-old brother butchered them and fed them to the family's dogs. By the end of the month, both men were dead. The older brother was buried before tissue samples were collected, but the younger one was confirmed as Indonesia's 68th human H5N1 case and 52nd fatality.

Indonesia has the greatest number of human H5N1 fatalities, 57 as of 17 December, and the highest H5N1 fatality rate in the world. Each human case increases the risk the virus will adapt to human hosts, sparking the dreaded pandemic.

Most Indonesians who have contracted the disease are not commercial poultry farmers but, like the Bandung brothers, were exposed while doing routine chores. "The key is how to make this type of person understand the danger of carrying dead chickens around," says Bayu Krisnamurthi, chief executive of the Indonesia National Committee for Avian Influenza Control and Pandemic Influenza Preparedness, which has launched a massive public education campaign.

James McGrane, team leader for the U.N. Food and Agriculture Organization's Avian Influenza Control Program in Indonesia, says it would be culturally and logistically impossible to suddenly eliminate backyard poultry or shift from live markets to a centralized slaughter system. The country has an estimated 300 million backyard poultry and 13,000 live poultry markets.

Participatory surveillance teams are advising small holders on how to handle carcasses and urging them to put their chickens in coops at night, two small ways to improve biosecurity (see main text). The national avian influenza committee is also considering requiring live poultry markets to close periodically for cleaning and disinfecting. But passing the needed legislation will take time.

Until such measures are in place, the best hope of averting a global pandemic is rapid response. This involves monitoring to detect when a deadly virus has started spreading among humans, followed by quarantines, widespread administration of antiviral drugs, and other measures.

In early October, Krisnamurthi's committee held the first of a planned series of simulations, a "tabletop" exercise to establish government responsibilities and lines of communication for pandemic response. Eventually, the government will hold a full-scale drill involving the army and other supporting personnel.

So far, scenarios for containing a budding pandemic presume that the deadly virus would emerge in a rural area. But "what if it's Jakarta?" Krisnamurthi asks, shaking his head. "It is definitely fair to say that … a very densely populated urban area … is a more difficult situation than a sparsely populated rural village," says World Health Organization virologist Keiji Fukuda.

Meanwhile, Indonesian and international epidemiologists are trying to understand why the fatality rate there is so high—75% as compared with about 67% in China and Thailand and 45% in Vietnam. According to Triono Soendoro, director general of the National Institute of Health Research and Development in Jakarta, epidemiologists are searching hospital records for retrospective bird flu cases. There are also plans to screen poultry farmers and cullers for antibodies indicating previous exposure.

Another puzzle is why one-third of Indonesia's human cases have come in clusters of blood relatives, like the brothers from Bandung. In some clusters, relatives by marriage had similar exposures but did not contract the disease. Genetic susceptibility could be involved, says Fukuda, as could different cultural patterns or a change in the virus. Unfortunately, Indonesia is likely to be a laboratory for human H5N1 infections for some time to come. **–D.N.**

Follow the trail

Those alarming results persuaded Indonesian authorities and international experts to push for a rapid expansion of participatory epidemiology. Even though coverage of the country is still spotty, the data being accumulated are providing clues to what keeps the virus in circulation.

For example, says Roeder, epidemiologists noted a curious pattern of outbreaks occurring after vaccination teams visited villages. They concluded that the teams were likely carrying the virus on contaminated clothing and vehicles and infecting the birds they vaccinated, which died before they developed immunity. Some villagers had recognized this pattern early on and started resisting vaccination, much to the puzzlement of authorities, who didn't make the connection. To cut such risks, the response teams are now training individuals within each community in vaccination.

Participatory response is an equally important part of the program. Mariner says that until recently the standard response was for government vets to cull all poultry in a broad swath around the villages where infected birds were found and then vaccinate widely. Local officials often feel that a show of force is politically necessary when a human case has turned up. This causes resentment among small holders, who may correctly believe that their birds have not been exposed to the virus. Telling small holders to cull their chickens "is like telling Americans to kill their dogs,"

CREDIT: AP

Mariner says. Delays in compensation exacerbate the ill feelings.

Instead, the participatory approach is to involve villagers in decisions—ideally, to cull all poultry directly exposed to infected birds, with immediate compensation, and then vaccinate other birds in the vicinity. Mariner says that even small holders can be convinced of the need to cull birds that have been directly exposed to H5N1-infected chickens.

Extending the reach of participatory response will require greater efforts to gain the understanding of local authorities. It will also require dependable funding for compensation and reliable supplies of vaccine. Mariner says some districts used up their yearly allotment of vaccine in a few months.

A bigger challenge is in proving that participatory epidemiology can begin to reduce the number of outbreaks. "I don't think we've had any impact on incidence [of outbreaks] so far," Mariner admits. Mariner and Jost are planning a new pilot program to measure which control and response measures have the most effect on outbreaks.

Growing a system

At the same time that Indonesia is verifying the effectiveness of participatory epidemiology, the country, with FAO support and financing from the United States, Australia, and Japan, is planning to extend the program to all of Java and Bali and parts of Sumatra by next May. That will require additional funding as well as more vets, or vet substitutes.

Roeder says he is confident that participatory epidemiology will have an impact. But the surveillance and response teams are just one link in an animal health infrastructure that should stretch from basic labs capable of tracking changes in the virus and verifying the efficacy of vaccines to better oversight of commercial poultry operations to a consistent response coordinated among all levels of government. The Ministry of Agriculture outlined these goals in a strategic plan a year ago. But financial support has been slow in coming.

Until that help comes, no one is willing to bet on when Indonesia will bring avian influenza under control. "This has gone past being an emergency program; we're in for the long haul," says FAO's Weaver. Which means that Sobandi and Rahmet Hidayat and their colleagues are just at the beginning of a long journey down the dusty lanes of Indonesia's villages.

–DENNIS NORMILE

BIOMEDICINE

Puzzling Out the Pains in the Gut

Newly identified mutations and immune cells are clearing up the mysteries of inflammatory bowel diseases and suggesting novel drug targets

CREDIT: ANATOMICAL TRAVELOGUE/PHOTO RESEARCHERS INC.

In an era when people often seem obsessed with maintaining a germ-free environment, it may come as a shock to many that the human intestinal tract is home to an estimated 100 trillion bacteria—microbes that help keep us healthy by producing certain nutrients such as vitamin K and fending off pathogenic bacteria that might otherwise colonize the gut. But even though we need our intestinal inhabitants, they present a conundrum, says immunologist Casey Weaver of the University of Alabama, Birmingham: "How can the innate immune system peacefully coexist with all these organisms?"

The innate system is the body's first line of defense against microbes, but there's increasing evidence, much of it genetic, that the coexistence Weaver finds so remarkable does fail on occasion. When that happens, the painful and debilitating conditions known as inflammatory bowel diseases (IBDs) may result. Together, the IBDs, which include Crohn's disease and ulcerative colitis, afflict 1 million people in the United States alone.

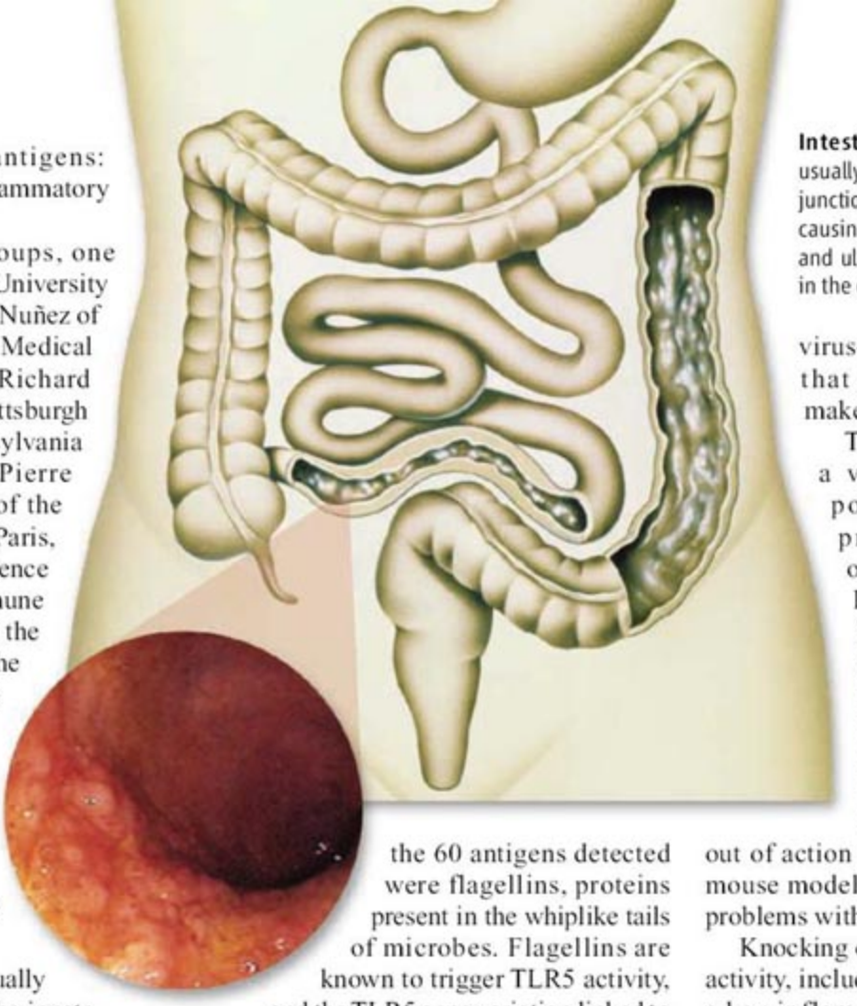
These intestinal problems do not arise solely in the innate immune system, however. The more microbe-specific adaptive immune response, which is activated by the innate system and works through T and B cells, can also accidentally prompt an attack on a person's guts. Apparently, the regulatory mechanisms that would otherwise keep both innate and adaptive immune responses in check are somehow disrupted. "The fundamental idea that many people have is that IBD is caused by an abnormal response to the normal gut flora," says Warren Strober of the National Institute of Allergy and Infectious Diseases in Bethesda, Maryland.

In one line of research that has received particular attention lately, immunologists, working mostly with animal models, have linked a newly discovered type of T cell known as the T_H17 cell to the inflammation underlying IBD, as well as to other autoimmune conditions. "The story's emerging that this [T-cell] lineage is a key factor in the progression of the disease," says immunologist Charles Elson of the University of Alabama, Birmingham.

The growing understanding of how the innate and adaptive immune systems can ravage the gut could lead to new treatments for IBDs. For example, researchers are turning up immune system molecules, including a cytokine that regulates T_H17-cell activity, that may be targets for more specific drugs. Current IBD treatments, such as corticosteroids, often rely on suppressing the whole immune system and therefore leave patients extremely susceptible to infection. Immunologists hope that drugs that block just the specific immune activity underlying the gut's abnormal inflammation will have fewer such side effects.

Trouble on the frontline

Studies of mice that spontaneously develop IBD provide one line of evidence that abnormal immune responses to microbial gut flora are a source of trouble. The mice, which develop persistent colitis in a normal environment, do not do so if reared in a germ-free environment, presumably because they have lost the ability to regulate immune responses to their intestinal partners. The bacteria of the gut, Strober says,



Intestinal problems. Crohn's disease usually affects the small intestine near its junction with the colon (*center cutaway*), causing the lining to be inflamed (*inset*) and ulcerated. Ulcerative colitis develops in the colon (*right cutaway*).

"are in a sense like self-antigens: always there" to keep an inflammatory response going.

Five years ago, two groups, one including Judy Cho of Yale University School of Medicine, Gabriel Nuñez of the University of Michigan Medical School in Ann Arbor, and Richard Duerr of the University of Pittsburgh School of Medicine in Pennsylvania and the other led by Jean-Pierre Hugot and Gilles Thomas of the Foundation Jean Dausset in Paris, provided the first direct evidence that an abnormal innate immune response to gut bacteria is at the center of some IBD cases. The researchers linked mutations in a gene then called *NOD2* (and since renamed *CARD15*) to increased susceptibility to Crohn's disease. Estimates are that mutations in this gene account for up to 15% of Crohn's cases.

The gene's protein, still usually called NOD2, is a sensor in the innate immune system that detects a common component of bacterial cell walls. On binding that component, it triggers inflammation and other responses. The mutations detected in the *CARD15* gene should result in production of a protein deficient in the ability to recognize and interact with its bacterial targets. Researchers have proposed several ways in which this apparent loss of the protein's function would result in disregulation of immune responses to intestinal bacteria and the persistent inflammation of Crohn's, but the issue remains unresolved.

Last year, a team including Cho, Duerr, and Andrew Gewirtz of Emory University in Atlanta, Georgia, linked a variant of another innate immune response gene to decreased, rather than increased, susceptibility to Crohn's disease, although only in Jewish patients. The gene in question encodes Toll-like receptor 5 (TLR5), another bacterial sensor protein. The discovery ties in with another recent development in IBD research: identification of what appears to be an important bacterial trigger for the intestinal inflammation.

About 2 years ago, Elson's team, in collaboration with that of Robert Hershberg at Corixa Corp. in Seattle, Washington, screened intestinal bacteria from mice with IBD for protein antigens that could contribute to development of intestinal inflammation. They found that about one-quarter of the 60 antigens detected were flagellins, proteins present in the whiplike tails of microbes. Flagellins are known to trigger TLR5 activity, and the TLR5 gene variation linked to decreased susceptibility for Crohn's disease should inactivate the receptor, thereby decreasing innate responses to the proteins. "By activating TLR5, the flagellins can drive inflammation directly through this immune pathway," Gewirtz says.

T_H17 cells debut

Illustrating the link between the innate and adaptive immune response, work by Hershberg, Elson, and their colleagues has revealed that several types of mice that develop IBD—and also some human Crohn's patients—have high levels of antibodies to the flagellins. Direct evidence that the bacterial proteins are involved in intestinal inflammation came when the researchers induced colitis in healthy mice by injecting them with flagellin-specific T cells from colitic animals. The T cells presumably contained so-called helper cells that could spark the production of antiflagellin antibodies by B cells. The flagellins probably aren't the only bacterial proteins able to promote colitis, however.

Over the years, immunologists have developed several lines of evidence linking T cells to autoimmune inflammation, including IBD. The early work may have led researchers astray, however. At the time, immunologists had recognized two distinct lineages of T helper (T_H) cells: T_H1 cells with functions including the destruction of virus-infected cells and T_H2 cells that cooperate with B cells to make antibodies.

The evidence, obtained with a variety of mouse models, pointed to T_H1 cells as the primary culprits in IBD and other autoimmune conditions. Researchers found large numbers of the cells in various inflamed tissues, for example. In addition, differentiation of T_H1 cells depends on a signaling protein called interleukin-12 (IL-12), and antibodies that take IL-12 out of action prevent inflammation in the mouse models. Even so, there were some problems with the T_H1 hypothesis.

Knocking out genes needed for T_H1 cell activity, including the gene for interferon γ, a key inflammation-stimulating protein made by the cells, should have prevented autoimmune inflammation from developing in the various mouse models. But in some models it didn't. If anything, the animals got even worse inflammation. "People were trying to put square pegs in the round hole of the T_H1-T_H2 paradigm," is how Weaver describes the situation.

The confusion began to clear about 6 years ago when Robert Kastelein and colleagues at Schering-Plough Biopharma in Palo Alto, California, discovered a new relative of IL-12, called IL-23. Each of these cytokines consists of two protein subunits. One, designated p40, is common to both, whereas the other—p35 in IL-12 and p19 in IL-23—differs. As it turns out, the antibody used in the experiments implicating T_H1 cells in autoimmunity reacts with the common subunit p40—and thus would block the activity of both IL-12 and IL-23. "That literature had to be completely reinterpreted," Kastelein says.

Kastelein, Daniel Cua, also at Schering-Plough Biopharma, and their colleagues went on to confirm that IL-23, not IL-12, is key to autoimmunity in mice. Working with two models of autoimmunity, one a brain inflammation similar to human multiple sclerosis and the other arthritis induced by collagen injections, the researchers compared the effects of knocking out the gene for IL-12, the one for IL-23, or both. Animals

without IL-12 still developed inflammation, whereas those without IL-23 did not.

In publications during the past year, the Schering-Plough team and another group led by Fiona Powrie of the University of Oxford in the United Kingdom have both shown that IL-23 is also needed for IBD development in mice. "IL-12–lacking animals go on to develop colitis," Powrie says, "but not those lacking IL-23."

Identification of this new interleukin led in turn to the discovery of T_H17 cells. About 2 years ago, Kastelein, Cua, and their colleagues reported that the cytokine promotes the development of a population of T cells very different from T_H1 cells. For example, these IL-23–responsive cells produce a different suite of cytokines, including the previously identified pro-inflammatory cytokine IL-17—hence the name T_H17 cells.

Weaver's team and also that of Chen Dong at M. D. Anderson Cancer Center in Houston, Texas, have further shown that the development of T_H17 cells is independent from that of T_H1 cells, requiring a completely different set of cytokines and other regulatory molecules. Researchers also confirmed that this T-cell lineage can promote autoimmune inflammation by showing that injection of the cells into mice induces brain inflammation.

The IL-23–T_H17 connection doesn't tell a complete story, however. In the mouse IBD model studied by Powrie and her colleagues, IL-23 can apparently induce intestinal inflammation independently of T_H17 cells: The animals in question lacked all T cells. The Oxford team's results suggest that the cytokine is instead working through the innate immune system, particularly via dendritic cells, which serve on the system's frontline as antigen detectors. "That's not to say that [IL-23] doesn't contribute to adaptive T-cell responses," Powrie says, noting that there are several modes of tissue inflammation that the cytokine could drive.

Inflamed. Interleukin-17 (a predicted structure *above*) is an inflammatory cytokine produced by T_H17 cells that may be major players in IBDs.

A link to human IBD

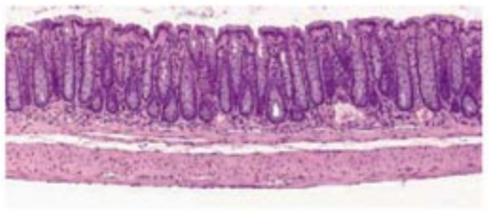
Although virtually all the work so far on IL-23 and T_H17 cells has been in mice, there is some evidence that the cytokine is involved in a human IBD. As reported online in *Science* on 26 October, a multi-institutional team led by Duerr and Cho performed a genomewide survey looking for gene variations called single-nucleotide polymorphisms (SNPs) associated with the development of Crohn's disease (*Science*, 1 December, pp. 1403 and 1461). The strongest association uncovered by the survey was a SNP located in a gene that encodes a receptor for a familiar cytokine—none other than IL-23. "The human genetics provides a very strong confirmation of what the immunologists have found," Cho notes.

That particular SNP, which causes a substitution of the amino acid glutamine for arginine at one location in the receptor protein, is apparently protective, decreasing the chances that people carrying it will come down with Crohn's. But the researchers also found a couple of weaker associations with SNPs in the receptor gene that increase risk. Cho and her colleagues are currently trying to understand how the gene variations influence development of Crohn's. "We think [the protective SNP] might impair IL-23 function, but that's a guess," she says.

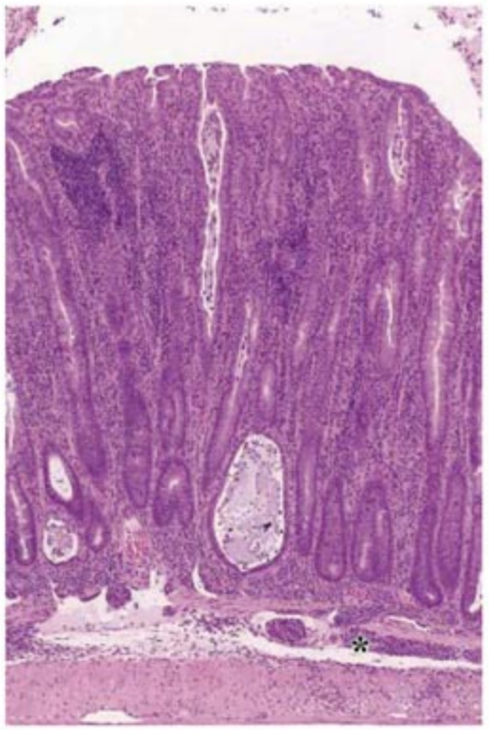
Results of an early clinical trial also lend support—although it's not unequivocal—to the idea that IL-23 plays a causative role in human IBD. The trial, involving a total of 79 people, was designed when T_H1 cells were the leading IBD culprits, and its aim was to test the safety and efficacy of an antibody directed against IL-12 in patients with Crohn's disease—the expectation being that the antibody would thwart stimulation of T_H1 cells. But that antibody, like those used in the animal studies, blocks the p40 subunit and so should also inhibit the activity of IL-23.

The study showed the antibody to have few side effects, mostly soreness at the antibody injection sites. And although not all dosing regimens showed signs of efficacy, 12 of the 16 patients who received the higher antibody dose given once per week for seven straight weeks had decreased symptoms compared to only two of the eight controls. (The results appeared in the 11 November 2004 issue of the *New England Journal of Medicine*.) "The antibody produced a very good response in the patients," says Strober, who was a member of the clinical research group.

Because the antibody reacts with both IL-12 and IL-23, it's not possible to say that only IL-23 is at fault in Crohn's. Indeed, Strober and his colleagues have found



Colitis trigger. In a mouse strain that spontaneously develops colitis, the colon lining is thickened and inflamed (*below*), but when that strain can't make IL-23, the colon appears normal (*above*).



that the concentrations of both cytokines are increased in human Crohn's patients and that treatment with the anti-p40 antibody causes the concentrations of each to go down. "The question [of which cytokine is more important] will not be answered until treatment is tried with an anti-p19 antibody directed at IL-23," Strober says. Such antibodies are reportedly under development at Schering-Plough.

Given the complexity of the IBDs, researchers expect that additional molecules and their genes are going to influence susceptibility to the painful conditions. The discoveries so far "are just the tip of the iceberg," Duerr says. Other candidates include two genes called *OCTN1* and *OCTN2*, which encode ion transporters. Geneticists are also looking at additional chromosomal sites linked to the diseases.

"Over the next months to a year, many of the low-hanging fruit will be harvested," Duerr predicts. "Then we can start doing analysis for gene-gene and gene-environment interactions and start putting all the pieces together."

–JEAN MARX

CREDITS (TOP TO BOTTOM): S. HYMOWITZ AND M. STAROVASNIK; DAN CUA *ET AL.*; D. YEN *ET AL.*, *THE JOURNAL OF CLINICAL INVESTIGATION* **116**, 1310 (2006)

Could Mother Nature Give the Warming Arctic a Reprieve?

The high Arctic appears to be on a slippery slope headed for a total meltdown. Year after year of shrinking ice and hungry polar bears seem to foretell immediate greenhouse oblivion. Now, though, some climate scientists say the region—the poster child of global-warming activists—may get a temporary reprieve.

"The [recent] warming has to have a natural component," explains Arctic researcher James Overland of the U.S. National Oceanic and Atmospheric Administration's Pacific Marine Environmental Laboratory (PMEL) in Seattle, Washington. When that natural trend inevitably reverses, he says, "it's very likely we'll see a slowing of the warming rate in the Arctic for a while." The region could even cool, until the next natural swing to the warm side again reinforces the greenhouse. A pause in warming could take the wind out of the sails of global-warming activists and bolster climate contrarians who say it's all just Mother Nature fiddling with climate.

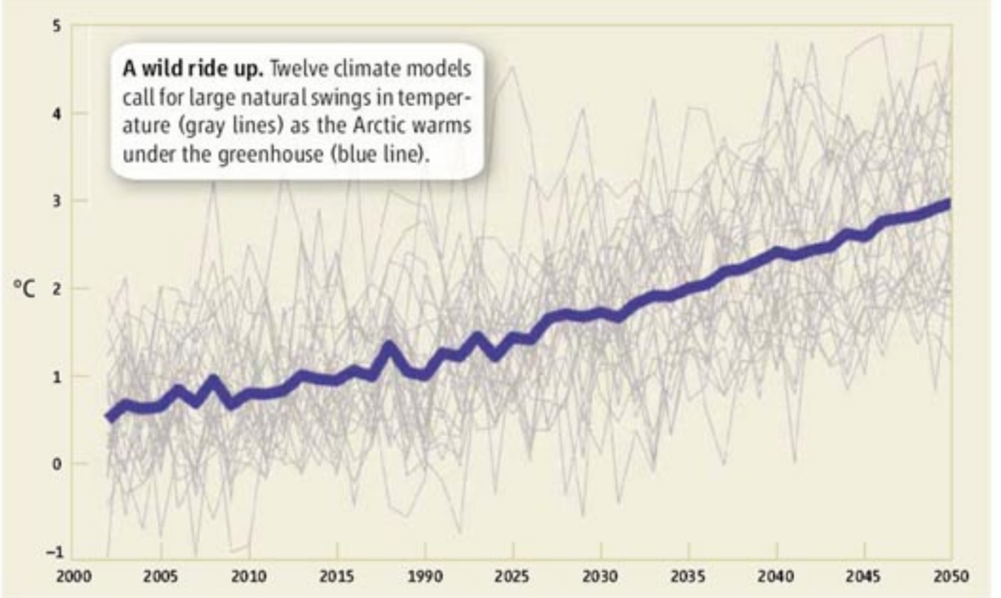
At the meeting, Overland and his University of Washington colleague Muyin Wang presented climate-modeling results that point to a major role for the Arctic's natural climate swings. First, they screened 22 climate models for how well they mimicked the region's ups and downs when human-produced "anthropogenic" changes to atmospheric gases were still small. Temperatures in the Arctic swung sharply from decade to decade in the 20th century; 10 of the models couldn't produce that much natural variability and were dropped.

Then Overland and Wang added the human factor. They evaluated the remaining 12 models run under the slowly rising greenhouse gas levels of the early and mid–20th century, the more sharply rising greenhouse forcing of the late 20th century, and then a best guess of how greenhouse gases will build in the 21st century. The models produced midcentury warmings reasonably similar to an actual Arctic warming that spanned the years from 1930 to 1950, Overland and Wang reported, but the timing of the model warmings varied by many years. If the still-weak greenhouse were driving them, they should have all occurred at the same time, so Overland and Wang take the real-world warming to have been natural variability.

Toward the end of the 20th century, however, all the models produced a substantial warming at the same time. That, they believe, is the greenhouse kicking in. But "the air temperature and loss of sea ice [in the Arctic] is occurring faster than global climate models would predict," says Overland. "I'm interpreting that as a fairly strong natural variability signal on top of long-term anthropogenic change." If he's right, the strong warming and accelerating ice loss of the past 5 or 10 years are the product of combined natural and humanmade warmings.

In the models' futures, natural climate variability still looms large. "It will be a bumpy road," says Overland. "It's very likely we could have a 5-year period of colder temperatures, and people could say, 'Aha, we don't have global warming.' " But the next natural swing to the warm side would once again add to greenhouse warming and would likely "send us to a new place," he says—a place with far less ice and even hungrier polar bears. "Jim has a very valid message there," says Arctic researcher John Walsh of the University of Alaska, Fairbanks, but a tricky one to get across. The average citizen may have trouble grasping greenhouse warming that—even only regionally and for a few years—simply goes away.

Arctic Temperature Anomaly Relative to 1980–1999



CREDIT: J. OVERLAND AND M. WANG

Weather Forecasting Way Out There

These days, forecasting rain or shine tomorrow means running huge number-crunching simulations on some of the biggest computers around. Meanwhile, "space weather" forecasters are still not much beyond the "Red sky at night, sailor's delight" stage of their science. Such rule-of-thumb methods have proved handy for predicting the geomagnetic storms that roil Earth's magnetosphere, fire up the aurora, and endanger satellites. At the meeting, however, researchers reported new progress but warned that empirical forecasting is approaching its practical limits. Scientists will have to answer some basic questions to make further progress in the field.

Space physicist Patricia Reiff of Rice University in Houston, Texas, and colleagues have been forecasting geomagnetic storms by using low-flying satellites to measure the speed of charged particles wafting through Earth's upper ionosphere. "We haven't missed a major storm for 2 years," Reiff boasts. The method works, she says, because the measurements reveal how the next region out—the teardrop-shaped, plasma-filled magnetosphere—responds to the solar wind, which ultimately drives magnetic storms. A few hours' worth of measurements can predict magnetospheric conditions during the next few hours. The Rice scheme also learns from past storm behavior so that the final forecast is not thrown off by the commonplace "quiet before the storm," Reiff says.

Snapshots From The Meeting >>

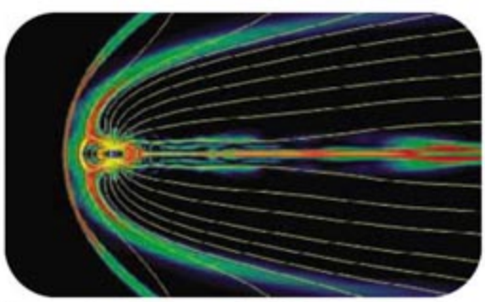
No toys. V. Ramanathan sent his fleet into the wild, cloudy yonder to untangle clouds' role in climate change.

Taking flight. After a couple of decades of development, crewless airplanes have arrived in atmospheric science. A trio of autonomous unmanned aerial vehicles (AUAVs) served as sensor platforms over the Indian Ocean last March during the Maldives Air Campaign, reported atmospheric scientist V. Ramanathan of Scripps Institution of Oceanography in San Diego, California. Under computer control, one of the 20-kilogram AUAVs would fly its 5-kilogram payload of miniaturized instruments above a cloud to measure incoming sunlight. Another would fly through the same cloud directly beneath the first to measure the properties of cloud particles and sunlight's interaction with them. And a third would fly beneath the cloud to characterize the pollutant particles from India and the Arabian Peninsula rising into the cloud. Such simultaneous in situ observations should help researchers solve a knotty problem of global warming: Are pollutant hazes masking some greenhouse warming by altering clouds?

A nastier early mars. When the Opportunity rover sent back signs of water early in martian history, the usual descriptor was "shallow salty seas." Sounded nice and cozy for any early martian life. But at a press conference at the meeting, rover science team leader Steven Squyres of Cornell University made a point of spelling out the team's best current understanding of early Mars, which is much less encouraging. "At the surface, this was primarily an arid environment," he said. Only occasionally, here and there, would puddles of salty, acidic groundwater form between dunes of salt sand. As the team's latest paper puts it, "dominantly arid, acidic, and oxidizing" environmental conditions would have posed "significant challenges to the origin of life." **–R.A.K.**

So far, the Rice forecast has been 98% accurate, Reiff reported at the meeting, with very few false alarms. It reliably gives a few hours' warning to satellite operators concerned about damage to sensitive electronics and to managers of vulnerable power grids on the ground. However, "we're reaching the limits of empirical forecasting," says Reiff. The next step must be full-blown simulations of space weather like those for earthly weather, preferably starting back at the sun, but "there's still some physics we have to learn," says Reiff.

For example, researchers still lack a clear picture of how the solar wind drives space weather. At the meeting, space physicist Patrick Newell of Johns Hopkins University's Applied Physics Laboratory in Laurel, Maryland, described how he and colleagues tried to nail down which aspects of the wind are most important for determining magnetospheric conditions such as the power of the aurora. After testing scores of formulas, they found that the best one involved just three properties of the solar wind. "The amazing thing is he fit 10 different data sets" with a single formula, says space physicist George Siscoe of Boston University.



CREDITS (TOP TO BOTTOM): V. RAMANATHAN; K. SIEBERT/MISSION RESEARCH CORP.

Lingering mystery. Earth's magnetosphere (simulated here) can still be hard to predict.

Newell thinks the correlation shows that the key to space weather is the rate at which the solar wind's magnetic field lines couple to Earth's magnetic field lines—an area that Siscoe and others pioneered. But Siscoe disagrees with that interpretation. Newell's formula is "telling us something about the coupling process that we don't understand," he says.

The Earthquake That Will Eat Tokyo

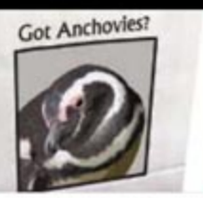
Denizens of the megalopolis of Tokyo are finally emerging from the threat of their own Big One. The megaquake that last struck offshore in 1923, killing 105,000 people, is not likely to return for many, many decades, researchers in a joint U.S.-Japan study reported at the meeting. But the same study finds that a far more immediate threat—including possible losses totaling $1 trillion—lies right beneath Tokyo and surrounding cities.

One-quarter of Japan's 127 million people live in and around Tokyo on the Kanto Plain. Unfortunately for them, not one but two tectonic plates converge on Japan from the east and dive beneath the edge of the Eurasian Plate and Tokyo. Sliding plates sticking and then snapping free produced quakes of about magnitude 8 in 1923 as well as in 1703.

How frequently do such quakes strike offshore Japan? A 20-member group asked that question at the meeting. The group was headed by seismologists Ross Stein of the U.S. Geological Survey in Menlo Park, California, and Shinji Toda of the Active Fault Research Center in Tsukuba, Japan, and funded in large part by the insurance giant Swiss Re. Each of those great earthquakes lifted the shoreline by a meter or more. That rise created wave-cut terraces perched above present-day beaches, preserving 7000 years of quake history in the terraces. By dating them, "Team Tokyo" researchers found that the last 17 quakes struck about every 400 years on average with surprising regularity. The probability of the next great quake striking in the next 30 years is then just 0.5%, the group reported.

Tokyo didn't get off so easy when Team Tokyo tackled the frequency of smaller quakes beneath the Kanto Plain. To judge by the frequency of earthquakes striking right beneath greater Tokyo, large quakes like the magnitude 7.3 shock of 1855 have about a 20% chance of occurring in an average 30-year period. Combining the two results, the chances of severe shaking in and around Tokyo are about 30% for the next 30 years, the group found, due almost entirely to the threat from beneath the city.

Using seismic records of 300,000 earthquakes in the area, the group believes it has pinned down the source of most of Tokyo's moderate but close-in quakes: a 25-kilometer-thick chunk of the Pacific Plate broken off and stuck between the three plates beneath Tokyo. Until that jam clears in the geologic future, residents of greater Tokyo will live under the threat of a trillion-dollar catastrophe rising from beneath their feet. At the meeting, seismologist David Jackson of the University of California, Los Angeles, raised the possibility that the threat is even larger than that. The great offshore quakes may not be as periodic as Team Tokyo would have them, he warned. The next one might misbehave and come sooner than expected. **–RICHARD A. KERR**



LETTERS

edited by Etta Kavanagh

The Israeli-Palestinian Science Organization

ON THE OCCASION OF THE NOVEMBER 2006 ANNUAL MEETING OF THE ISRAELI-PALESTINIAN Science Organization (IPSO), we, the members of IPSO's International Scientific Council, noted with considerable satisfaction the receipt of 71 proposals for joint scientific research between Palestinian and Israeli scientists, engineers, health professionals, and scholars who wish to work together.

In addition to its goal of promoting and funding joint research, IPSO encourages and supports quality education of Palestinian students and researchers—an essential element in securing a stable and economically viable society. Because of its desire to create a science-based bridge of good will, cooperation, and dialogue, IPSO joins Israeli university rectors and professors in opposing the ban that prohibits residents from the Palestinian Authority (PA) areas from entering Israel to study or to reach educational institutions in PA areas. We also call on the Israeli security authorities to allow, on an individual basis, academically qualified students to study in Israel.

> "We must actively promote favorable conditions for all to meet and work together in a safe, equitable, and productive environment."
>
> —Wiesel *et al.*

IPSO endorses the 31 October 2006 statement by the Council of the Israel Academy of Sciences and Humanities that calls on the Israeli government to "refrain from instituting any policy that hinders any group of scientists or academics, whether Palestinian or otherwise, from properly discharging their academic responsibilities."

Lastly, we urge the international scientific and scholarly community to support IPSO's goals of promoting high-quality research, advancing training in all areas of science and learning, and forthrightly opposing obstructions to academic freedom worldwide, including boycotts, moratoria, and arbitrary or sweeping access restrictions on students and academics to universities and research institutions. We must actively promote favorable conditions for all to meet and work together in a safe, equitable, and productive environment.

TORSTEN WIESEL,[1*] PETER AGRE,[2] KENNETH J. ARROW,[3] MICHAEL ATIYAH,[4] EDOUARD BRÉZIN,[5] FAOUZIA FARIDA CHARFI,[6] CLAUDE COHEN-TANNOUDJI,[7] ABDALLAH DAAR,[8] FRANCOIS JACOB,[9] DANIEL KAHNEMAN,[10] YUAN TSEH LEE,[11] IDA NICOLAISEN,[12] SARI NUSSEIBEH,[13] HARALD REUTER,[14] YOAV SHOHAM,[15] JOHN SULSTON,[16] MICHAEL WALZER,[17] MENAHEM YAARI[18**]**

[1]President Emeritus, Rockefeller University, New York, NY, USA. [2]Duke University, Durham, NC 27710, USA. [3]Stanford University, Stanford, CA 94305, USA. [4]University of Edinburgh, Edinburgh, Scotland EH9 3JZ, UK. [5]École Normale Supérieure, F-75231 Paris Cedex 05, France. [6]Institut Préparatoire aux Études Scientifiques et Techniques, Tunis, Tunisia. [7]Collège de France, 75231 Paris Cedex 05, France. [8]University of Toronto, Toronto, ON M5G 1L7, Canada. [9]Institut Pasteur, 75724 Paris CX 15, France. [10]Princeton University, Princeton, NJ 08544, USA. [11]Academia Sinica, Taipei, Taiwan (Chinese Taipei). [12]Nordic Institute of Asian Studies, DK-2300 Copenhagen, Denmark. [13]Al-Quds University, Beit Hanina, Jerusalem, Israel. [14]University of Bern, CH-3010 Bern, Switzerland. [15]Stanford University, Stanford, CA 94304, USA. [16]Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK. [17]Institute for Advanced Study, Princeton, NJ 08540, USA. [18]Jerusalem, Israel.

*Chair, IPSO

**Deputy Chairs, IPSO

Another Nail in Which Coffin?

A RECENT REPORT "VOLCANISM IN RESPONSE to plate flexure" by N. Hirano *et al.* (8 Sept., p. 1426) on small-volume volcanoes located far from plate boundaries and the related Perspective by M. McNutt "Another nail in the plume coffin?" (8 Sept., p. 1394), which casts this contribution in the context of the ongoing plume controversy, have inspired us to comment. Hirano *et al.* begin with the claim that "[v]olcanism on Earth is known to occur in three tectonic settings: divergent plate boundaries…, convergent plate boundaries…, and hot spots," followed by "Without the presence of a hot spot, new volcanism is not anticipated." They go on to show that plate flexure can induce fracturing and small-volume volcanism, which is unlikely to be related to any "mantle plume." The tenor of McNutt's Perspective is that this constitutes a victory in the ongoing battle against the hypothesis that volcanoes at hot spots are caused by jets of hot material (plumes) rising from the deep mantle ("It is thus with much kicking, dragging, and screaming that geoscientists are being brought to the realization that all might not be well with the concept of mantle plumes"). Actually, small-volume, within-plate volcanism isn't exactly news. There is a wealth of well-established knowledge about three types of within-plate, often alkalic volcanism that cannot directly be caused by deep mantle plumes: (i) tens of thousands of small volcanic seamounts (*1*, *2*); (ii) the 1600-km-long chain of oceanic and continental volcanoes known as the Cameroon Line, which shows no detectable time progression of eruptions and has long been discussed as a "hot line," not a plume (*3*); and (iii) volcanism associated with continental (and oceanic) rifts not related to plate boundaries. The discovery of another line of small alkalic seamounts not caused by a hot spot or a plume is neither new nor surprising. And the suggestion that this constitutes a "nail in the plume coffin" merely obfuscates the plume debate.

The basic rules of geochemistry tell us that high enrichments of incompatible trace elements found in alkali basalts require low degrees of melting; such melts are hardly a surprise in regions of thick oceanic lithosphere. Whether the volcanism itself is triggered by a mantle plume or a fracture in the lithosphere is a separate question. The observation that fractures can and do trigger volcanism is not an argument against deep-mantle plumes, any more than confirmation of mantle plumes, for example, through seismic tomography (*4*), could be an argument against fracture-related volcanism.

The question of whether deep-mantle plumes exist is too important to our understanding of mantle dynamics to be addressed in this fashion. Instead, let's look at those volcanic features where a plume mechanism actually makes some geological sense and investigate those. We suspect that geochemistry will not deliver the silver bullet for proving or disproving plumes. Rather, we suspect that when the dust has settled over the mapping of plumes with seismic tomography, we will come to a consensus over the question of whether the Hawaiian hot spot, for example, is caused by a plume. Evidence from small seamounts seems completely irrelevant to this debate.

So before we nail any more coffins, let's first be sure that there is a body to be buried.

ALBRECHT W. HOFMANN[1] AND STANLEY R. HART[2]

[1]Max Planck Institute for Chemistry, Postfach 3060, D-55020 Mainz, Germany. E-mail: hofmann@mpch-mainz.mpg.de. [2]Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA.

References

1. A. Zindler, H. Staudigel, R. Batiza, *Earth Planet. Sci. Lett.* **70**, 175 (1984).
2. R. Batiza, D. Vanko, *J. Geophys. Res.* **89**, 11235 (1984).
3. J. G. Fitton, H. M. Dunlop, *Earth Planet. Sci. Lett.* **72**, 23 (1985).
4. R. Montelli *et al.*, *Science* **303**, 338 (2004).

Response

WE FOUND YOUNG ALKALIC VOLCANOES, NOT older than 1 million years, on the edge of 135-million-year-old oceanic crust in the Northwest Pacific, where it is subducting into the Kuril and Japan trenches. These volcanoes are closely associated with the occurrence of extensional cracks in the flexed parts of the subducting Pacific Plate, allowing small amounts of partial melts to find their way to Earth's surface from the shallow asthenosphere. Because of their small volume, we named these volcanoes petit spots, which should not be confused with hot spot volcanism.

Hofmann and Hart fittingly note that McNutt's Perspective stretches her interpretation of this new type of within-plate volcanism too thin, by linking it to the heated debate on the existence of mantle plumes and the formation of the major hot spot trails, like Hawaii (*1*, *2*). We agree with Hofmann and Hart that the discussion of the petit spot model should be divorced from the ongoing mantle plume debate. Instead, we emphasize the unique tectonic settings in which the petit spot volcanoes are formed.

Other types of non–plume-related volcanism have been recognized to exist close to the midocean spreading centers, in continental rifts, and maybe in some seamount trails, like the Cameroon Line. However, none of these volcanoes were formed at locations far away from spreading centers, hot spots, or, more generally, areas of thermal upwelling. Whereas small off-axis seamounts (*3*–*5*) are easily explained by the faulting and thermal contraction of juvenile oceanic crust during seafloor spreading, the formation of young volcanoes on oceanic crust older than 100 million years is rather uncommon and remained undiscovered until we recognized the petit spots on the oldest part of the Pacific Plate. These volcanoes have geochemical signatures that are characterized by highly alkaline major element compositions, highly enriched incompatible elements, and degassed noble gas isotope ratios. These characteristics suggest a low degree of partial melting in combination with an origin in the shallow upper mantle, about 95 km deep.

The petit spots thus should be accepted as a new type of within-plate volcanoes that are not fed by large-scale thermal upwellings or mantle plumes. However, it also is a rather uncommon type of volcanism, which only represents a minute fraction of the total volcanic output in the ocean basins. As Hofmann and Hart argue, this type of volcanism is entirely unrelated to the processes that may form the voluminous Hawaiian seamount trail. In fact, the petit spots can be entirely explained by the bending and cracking of the subducting Pacific Plate, which is a rather unique situation and may only be reserved for oceanic crust that is located close to a convergent plate boundary.

NAOTO HIRANO AND ANTHONY A. P. KOPPERS

Scripps Institution of Oceanography, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093–0225, USA. E-mail: nhirano@ucsd.edu; akoppers@ucsd.edu

References

1. G. R. Foulger, J. H. Natland, *Science* **300**, 921 (2003).
2. A. A. P. Koppers, H. Staudigel, *Science* **307**, 904 (2005).
3. A. Zindler, H. Staudigel, R. Batiza, *Earth Planet. Sci. Lett.* **70**, 175 (1984).
4. R. Batiza, Y. Niu, W. C. Zayac, *Geology* **18**, 1122 (1990).
5. D. T. Sandwell *et al.*, *J. Geophys. Res.* **100**, 15087 (1995).

Letters to the Editor

Letters (~300 words) discuss material published in *Science* in the previous 6 months or issues of general interest. They can be submitted through the Web (www.submit2science.org) or by regular mail (1200 New York Ave., NW, Washington, DC 20005, USA). Letters are not acknowledged upon receipt, nor are authors generally consulted before publication. Whether published in full or in part, letters are subject to editing for clarity and space.

Response

I AGREE COMPLETELY WITH HOFMANN AND Hart that a nonplume origin for young volcanoes seaward of the Japan trench cannot be used to argue that plumes do not exist anywhere. As they point out, Earth is rife with examples of nonplume volcanoes that form along suspected fractures away from plate boundaries. What is unusual in this particular case is that, in addition to the fracturing mechanism being well constrained rather than just "surmised," small volumes of melt penetrated a very thick plate above a region of large-scale downwelling, all features that are the antithesis of what is expected from plume theory. The existence or nonexistence of plume-type transport of heat and mass bears on the rheology of Earth's mantle, thermal and chemical layering in the interior, mixing rates of geochemical heterogeneities, the energy budget of the geodynamo, and other properties that are difficult to assess deep within this dynamic planet. Indeed, high-resolution seismic imaging holds the best hope for settling this debate. However, further gains in resolution at the scale needed to resolve plumes require filling in the very large gaps in network coverage in the ocean basins with seismic receivers, one of the goals of the Ocean Observatories Initiatives of the U.S. National Science Foundation. Recently, concerns over the cost of installing and maintaining deep-water open-ocean seismic observatories have led to a reduction in the plans for filling in these gaps. Sadly, unless some new source of funding or more affordable technology can be found, 10 years from now, our "patient" might still be lying on life support.

MARCIA MCNUTT

Monterey Bay Aquarium Research Institute, 7700 Sandholdt Road, Moss Landing, CA 95039, USA. E-mail: mcnutt@mbari.org

Chemistry Nobel Rich in Structure

THE 2006 NOBEL PRIZE IN CHEMISTRY, awarded to Roger Kornberg for the structure and understanding of RNA polymerase ("Solo winner detailed path from DNA to RNA,"

R. F. Service, News of the Week, 13 Oct., p. 236) (*1*), marks the latest in a long line of Nobel Prizes awarded in the area of macromolecular structure analysis.

It is interesting to consider how the selection of Nobel Prizes over the past few decades reflects our fascination with the structure of biomolecules.

In particular, there were 12 Nobel Prizes in chemistry and physiology or medicine awarded for work in this field from 1956 to 2006 (table S1) (*2*). Almost one in four chemistry prizes since 1956 have been for structure work, and in the last decade, fully half have dealt with work related to macromolecular structure.

Because many of these prizes were awarded for fairly recent work, we can compare their subjects to those in the scientific literature as represented in publication databases.

We examined the relative abundance of papers dealing with "protein conformation" and its subset, "crystallography," by measuring the number of records matching these Medical Subject Heading (MeSH) terms in PubMed for each year from 1970 to the present (table S2) (*2*). (We assumed that the sum total of PubMed publications bearing the chemistry MeSH term represent the corpus of work eligible for a Chemistry Nobel.) From 1970 to 2006, 4% of all chemistry publications dealt with crystallography, yet this subfield captured 19% of the available Nobel Prizes (table S3) (*2*). During the past decade, crystallography papers represented 7% of all chemistry publications, but commanded 4 of 10 available prizes. Even the much broader category of protein conformation displays two-fold "Nobel enrichment" in both year ranges. Overall, the Nobel Prizes in chemistry are noticeably enriched for work in macromolecular structure determination.

Macromolecular structure determination is a potent tool to understand biological systems and periodically yields landmark results that impact the scientific community at large. It would also seem that the surest road to Stockholm is through a crystal tray.

MICHAEL SERINGHAUS AND MARK GERSTEIN

Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520, USA.

References

1. P. Cramer, D. A. Bushnell, R. D. Kornberg, *Science* **292**, 1863 (2001).
2. See Supporting Online Material available at www.sciencemag.org/cgi/content/full/315/5808/40/DC1.

CORRECTIONS AND CLARIFICATIONS

Random Samples: "Demise of a blimp" (3 Nov. 2006, p. 735). The subject of this item, the USS *Macon*, was misidentified as a blimp. It is a dirigible.

ANTHROPOLOGY

All Is Not Always Lost When the Center Does Not Hold

Kathleen D. Morrison

After Collapse
The Regeneration of Complex Societies
Glenn M. Schwartz and John J. Nichols, Eds.
University of Arizona Press, Tucson, AZ, 2006. 295 pp. $50. ISBN 97808165096.

It is perhaps not too surprising that in a time of widespread anxiety about global environmental change, the collapse of civilizations is a topic of intense interest. For example, building on scholarly work in archaeology from the last 20 years, Jared Diamond's recent bestseller, *Collapse* (*1*), merged an apocalyptic vision of environmental degradation with an upbeat lesson from the self-help literature, suggesting that societies have chosen to succeed or fail. Spectacular failures, real or imagined, certainly have broad popular appeal: Regional abandonments of settlements in the U.S. Southwest are typically represented as "mysterious disappearances" even though such shifts represented common and effective strategies for managing ecological challenges—and, of course, even though Pueblo peoples are still very much present in the region today.

Similarly, the Maya are famous primarily for having a complex political and social order involving monumental architecture, an order that failed spectacularly in the Terminal Classic period. Although the Classic Maya collapse involved both political change and large-scale depopulation, even there life went on: as Diane Chase and Arlen Chase describe in their contribution to *After Collapse*, Postclassic Mayan society restored Classic-period institutions of symbolic egalitarianism and shared rule while rejecting Terminal Classic strategies that more clearly marked personal inequalities. Neither Mayan civilization nor Mayan peoples disappeared, a long-term record of continuity that seems to be the norm rather than the exception, as the articles in this volume make clear.

Archaeologists and historians have long recognized the often-unstable nature of complex societies, using notions of cyclical "rise and fall" or organic metaphors of growth, maturity, and senescence to describe organizational changes in early states. Robert McCormick Adams and Norman Yoffee, both working on early states in the Middle East, formulated influential frameworks suggesting that oscillation between periods of centralization and urbanization and periods of ruralization and local rule was a normal feature of early complex polities (*2*, *3*). Thus, fluctuations, rather than the steadily increasing complexity posited by dominant models of cultural evolution, should be expected.

As Glenn Schwartz's introduction to *After Collapse* points out, however, not all of these phases have been equally well studied. Studies of state collapse and of the initial development of complex societies have continued to be counted among the big questions of archaeology. Why the regeneration of complex societies after episodes of collapse has not, to date, been a major focus of research can be attributed to an archaeological obsession with origins and in particular with "primary states," those six places where complex polities developed without prior organizational models. The diffusionary logic that the idea of the state was somehow a sufficient condition for the emergence of complex polities has been long discredited, yet for some reason archaeological disregard for so-called "secondary state formation" has continued. Not only do the vast majority of cases of state development fall under this rubric, but so do instances of regeneration after collapse. Hence, the reasons for underanalysis of this important process are, if not clear, at least explicable. What all this suggests is that the examples presented in *After Collapse* have the potential to inform on processes of state (re)formation more generally; addition of these important cases can only add to our understanding of state generation as well as regeneration.

Schwartz notes that the study of state regeneration is, in large part, a study of "dark ages," a term that, besides encoding value judgments developed under conditions of centralization, also refers to the paucity of textual information for periods after collapse. The negative valences of terms such as dark age and even collapse certainly reveal viewpoints firmly invested in text-based history (no period is darker than any other to an archaeologist) and in social hierarchy (what falls apart in a collapse are often structures of inequality). Archaeology, however, is well situated to address issues of change where texts disappear.

Here it is worth clarifying what contributors to this volume mean by collapse. As Schwartz enumerates, collapse "entails some or all of the following: the fragmentation of states into smaller political entities; the partial abandonment or complete desertion of urban centers, along with the loss or depletion of their centralizing functions; the breakdown of regional economic systems; and the failure of civilizational ideologies." Note that this definition refers only to the collapse of complex political structures and that death and destruction are conspicuously absent. Although the focus of *After Collapse* is decidedly on continuity and renewal, archaeological studies of collapse itself [e.g., (*4*)] have always recognized that civilizational traditions and peoples rarely disappear.

What, then, causes state regeneration and how does it proceed? Are, as Schwartz asks, such processes simply replays of earlier developmental episodes? Or are new strategies and trajectories involved? One might

From a Greek revival. The Temple of Poseidon (450–440 BC) at Sounion, Greece, was constructed early in the Classical Period, on the site of a slightly smaller, Archaic Period temple.

CREDIT: MARCO SIMONI/GETTY IMAGES

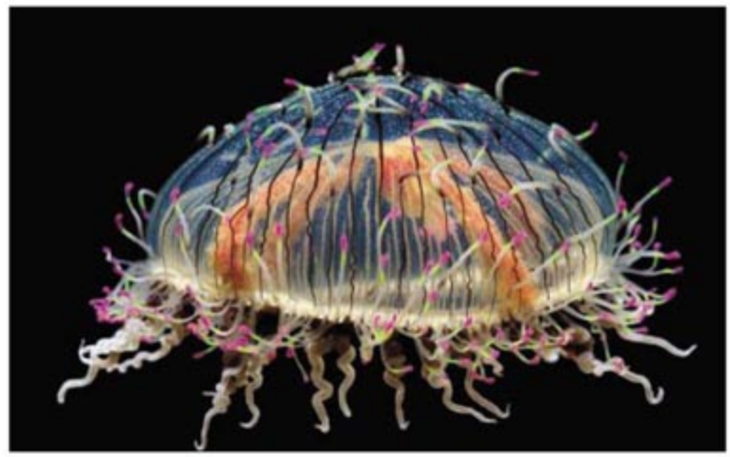
The reviewer is in the Department of Anthropology and the Center for International Studies, University of Chicago, 5828 South University Avenue, Chicago, IL 60637, USA. E-mail: morrison@uchicago.edu

BROWSING

Life.

A Journey Through Time. Frans Lanting, edited by Christine Eckstrom. Taschen, Cologne, 2006. 304 pp. $49.99, £29.99, €39.99. ISBN 9783822839942.

Although the subtitle might suggest a focus on fossils, here Lanting interprets the history of life on Earth through photographs of extant organisms and landscapes. Successive chapters highlight marine life (such as the flower hat jelly, *Olindias formosa*), shore dwellers, the spread of plants and vertebrates across the land, innovations of birds and flowering plants, and mammals. Aerial views in the concluding chapter mark how life shapes our planet. A slideshow featuring images from the book is at www.lifethroughtime.com, which also offers information about a multimedia orchestral performance (with music by Philip Glass) and a traveling exhibition.



CREDIT: © FRANS LANTING/WWW.LIFETHROUGHTIME.COM

think, given the popularity of climate- and resource-oriented explanations for collapse, that many scholars would place regeneration at the feet of climatic amelioration or environmental regeneration. However, with the exception of Ian Morris's careful exposition of the transitions from Mycenaean (Late Bronze Age) Greece through the Greek Dark Ages and on to the Classical Period, contributors to this volume have surprisingly little to say about environmental conditions. Perhaps this is because the Greek case, like the Classic Maya, is an example of what Bennet Bronson in this volume calls "genuine regeneration," not simply the shift of a political or economic center but a transformation of the entire system. Indeed, the differences between Classical and earlier periods are profound (with perhaps little more than the memory of a lost heroic age linking them)—a shift even more substantial than that seen in the Maya region, albeit one covering much longer periods of time.

Contributors analyzing other regions (including Egypt, Peru, Cambodia, and Bronze-Age Syria) favor either Bronson's "stimulus regeneration," state building explicitly based on a hazily understood model distant in space or time, or his "template regeneration," a revival process based on fully understood, well-recorded models, often states close to the revived polity in space and time. Although both of these terms evoke the language of early 20th-century diffusionism, they at least have the advantage of stressing the ways in which regenerating polities make use of existing models of and ideologies for systems of structured inequality.

While *After Collapse* also asks when regeneration might not appear, the volume presents only one such counterexample, Kenny Sims's analysis of the upper Moquegua Valley, Peru. There complex political forms failed to regenerate after the fall of the Tiwanaku and Wari empires. Sims argues that restriction of local residents to client status and, at best, mid-level positions within the Wari administration left them without the wherewithal to (re)generate a centralized state. The general enthusiasm for Bronson's memory and knowledge-oriented categories might reflect the selection of cases themselves, few of which are examples of more radical collapse, in which depopulation as well as deurbanization took place.

In many ways, both the strengths and weaknesses of *After Collapse* reflect larger trends in archaeology. Contributors carefully consider how, precisely, people managed (or failed) to regenerate a complex polity after a political collapse, including some interesting considerations of the ways in which collapse presented opportunities for previously marginal elites to become the central players in regenerated regimes. However, there is disappointingly little willingness to consider why, specifically, complex polities (re)emerged—to address the origins of the secondary state, to use the jargon. This is an important question, with implications for state formation in innumerable cases, well beyond the sample of collapsed polities. If, for example, as Lisa Cooper, building on the arguments of Yoffee and Adams, suggests of Bronze-Age Syria, village-based organization was actually more stable in the long term than urbanism, then perhaps the formation of a complex polity might itself constitute "collapse." Such a perspective, suggested only half-seriously in Yoffee's closing remarks, might actually be salutary in finally purging the discipline of its rise-and-fall thinking. This could bring us one step closer to using the great strength of archaeological research, its immense time depth, as a serious guide for contemporary considerations of the sustainability and continuity of civilizations in the face of rapidly changing natural and social conditions.

References

1. J. Diamond, *Collapse: How Societies Choose to Fail or Succeed* (Viking, New York, 2005); reviewed by T. Flannery, *Science* **307**, 45 (2005).
2. R. McC. Adams, *Proc. Am. Philos. Soc.* **122**, 239 (1978).
3. N. Yoffee, *Am. Antiq.* **44**, 5 (1979).
4. N. Yoffee, G. Cowgill, Eds., *The Collapse of Ancient States and Civilizations* (Univ. Arizona Press, Tucson, AZ, 1988).

10.1126/science.1133393

HISTORY OF SCIENCE

Alchemy and the Science of Matter

Pamela Smith

In 1661, when at the age of 34 Robert Boyle published *The Sceptical Chymist: Or Chymico-Physical Doubts & Paradoxes* (*1*), he was already a confirmed experimentalist and atomist, having the year before published *New Experiments Physico-Mechanical, Touching the Spring of the Air and its Effects* (*2*). Both Boyle's empiricism and his defense of the theory that matter is made up of minute corpuscles, or atoms, have always placed him firmly on the modern side of the divide of the Scientific Revolution, that epochal transformation from a geocentric to a heliocentric universe, from Aristotelian hylomorphism to atomism, and from Aristotle's four causes to the mechanical philosophy. It always seemed obvious to historians that Boyle must have drawn his atomism from physics because

The reviewer is at the Department of History, Columbia University, 1180 Amsterdam Avenue, MC 2527, New York, NY 10027, USA. E-mail: ps2270@columbia.edu

Hendrick Heerschop's *The Alchemist's Experiment Takes Fire* (oil on canvas, 1687).

alchemical theories of matter (such as the three-principle theory that all matter is composed of salt, sulfur, and mercury or the Aristotelian theory of the four elements bound together by form) gave no obvious source for atomism within chemistry itself. Boyle's well-known empiricism, however, was a different matter. Historians of chemistry, especially William Newman, have in recent years shown just how important the hands-on empirical traditions of alchemy and early modern chemical industries were in developing the attitudes and techniques that came to define the new experimental philosophy of the Scientific Revolution.

Newman's *Atoms and Alchemy* now demonstrates, through close study of some of the most prolix and dense chemical writers of the medieval and early modern periods, that Boyle's atomism too can be traced back to a tradition of alchemical experimentation and theorizing that began in the 13th century. Newman places the advent of corpuscular theories of matter in the alchemical writer Geber's *Sum of Perfection*. Geber is a partial Latin transliteration of Jābir ibn Hayyān, a possibly fabulous eighth-century author of many chemical works in Arabic. (Newman's doctoral dissertation was devoted to proving that the *Sum of Perfection* was actually the work of a 13th-century Latin author, probably Paul of Taranto.) Although Geber proposed a corpuscular theory of matter, he was no revolutionary but rather a dedicated follower of Aristotle, and his corpuscles were made up ultimately of the four elements that combined into atoms that joined to give rise to the two principles of sulfur and mercury, which in turn combined to form the metals.

The fact that Geber integrated the idea of corpuscles of matter into scholastic theory is not the only startling discovery. Equally surprising in terms of conventional history of science is that his theory was expressed in terms of his laboratory work in cupellation, cementation, and other assaying techniques. In other words, at this early stage in alchemistry, there is remarkable integration between theory and laboratory practice in the works of Geber and the known works of Paul of Taranto. These authors insisted that laboratory operations were capable of revealing the fundamental components of matter by analysis. Much of their investigations was devoted to reinterpreting Aristotle's concept of "mixture," and the alchemical authors rewrote the doctrine of *mixis* contained in Aristotle's *De generatione et corruptione*. Such reinterpretation aroused the ire of Thomas Erastus, a Heidelberg professor of medicine who, in a 1572 work seeking to debunk alchemists' claims that base metals could be transmuted into gold and silver, attacked this rewriting and the principle of understanding matter by chemical analysis in the laboratory.

Atoms and Alchemy
Chymistry and the Experimental Origins of the Scientific Revolution

by William R. Newman

University of Chicago Press, Chicago, 2006. 272 pp. $75, £47.50. ISBN 9780226576961. Paper, $30, £19. ISBN 9780226576978.

The 16th and early 17th centuries being what they were, Erastus's public attack in an extraordinarily long and dense Latin treatise in turn provoked pugnacious expostulations to the contrary in similarly weighty tomes. Among the responders, Andreas Libavius proposed, in his *Alchymia* (1606), an ecumenical combination of ancient atomism, Aristotelianism, and medieval alchemy. A Wittenberg professor of medicine, Daniel Sennert—puzzled in his laboratory investigations by what kind of mixture could give rise to the complete dissolution in nitric acid of silver in a gold alloy and then its complete recoverability through precipitation—employed Libavius's combination to propose his own version of Aristotelian atomism. Sennert, both a corpuscularian and a committed Aristotelian, retained the principles of matter and form but believed discrete atoms were little bundles of matter and form locked together into semi-permanent corpuscles.

Such a coexistence of atomism and Aristotelianism is one of the important points that emerges from Newman's close readings of works that have never before formed part of the familiar canon of the Scientific Revolution. In the last half of the 17th century, Robert Boyle took up Sennert's reasoning from his crucial nitric acid experiments as well as his view that individual atoms were Aristotelian *mixis*, and Boyle employed them to debunk the theoretical framework of Aristotelianism that Sennert was trying to defend. Boyle used Sennert's demonstrations of enduring corpuscles to develop and support his conception of a mechanical philosophy of matter. Newman persuasively shows that Boyle conceived of the mechanical philosophy in terms of a machine that can be visualized and known through observation of the interaction of its gears and springs, rather than as a philosophy dealing primarily with undifferentiated matter in motion. Boyle compared a chemical compound to a clock, in which the parts fit together to form an interactive structure, a mechanism. The goal of Boyle's chemical investigations was thus to demonstrate that the phenomena of the sensible world could be reduced to mechanical causes in this sense.

Newman's view that Boyle's mechanical philosophy was actually the outcome of a long tradition of alchemical experimentation and theorizing that sought to recast Aristotelian theories of mixture will be surprising to most historians and will impel a rewriting of the history of chemistry. Certainly such a view integrates alchemy firmly into the narrative of the rise of modern science. Newman offers additional support to the emerging consensus that not only was alchemy not a pseudoscience in the Middle Ages and the Renaissance, it was also central to the development of empiricism. In addition, *Atoms and Alchemy* shows that alchemy was absolutely crucial to two of the theoretical cornerstones of modern science, atomic theory and the mechanical philosophy.

References

1. http://oldsite.library.upenn.edu/etext/collections/science/boyle/chymist/.
2. Reprinted in M. Hunter, E. B. Davis, Eds., *The Works of Robert Boyle* (Pickering and Chatto, London, 1999).

10.1126/science.1131578

SUSTAINABILITY

Anchovy Fishery Threat to Patagonian Ecosystem

Elizabeth Skewgar,[1]* P. Dee Boersma,[1] Graham Harris,[2] Guillermo Caille[2]

Unsustainable anchovy harvest could disrupt ecosystem function, other fisheries, and ecotourism revenues.

CREDITS: JOEL JACOBS AND STOCKPHOTO

The Patagonian coast is famous for its charismatic megafauna—Magellanic penguins, southern right whales, southern elephant seals, and southern sea lions—but the small, less charismatic southwest Atlantic anchovy is a key trophic link in the ecosystem (*1*). Overfishing anchovy could disrupt energy flows in the southwest Atlantic ecosystem, harm other fisheries and wildlife, and damage the valuable ecotourism sector.

In 2003, Argentina's Federal Fisheries Council (CFP) approved a plan by the Province of Chubut for an experimental program to develop a small-scale trawler fishery for the "under-exploited" anchovy in provincial waters south of 41°S, partially as an alternative to the overfished hake (*2*). The plan notes the proximity of the Peninsula Valdés (a World Heritage Site) and the world's largest continental Magellanic penguin colony at Punta Tombo, but has no specific mechanisms to quantify the fishery's effect on the fish and wildlife species that depend on anchovy. In both 2004 and 2005, Argentine catches exceeded 30,000 tons of anchovy for the first time in 30 years (*3*).

Rising global demand for fish meal could fuel unsustainable anchovy fishery expansion on the Patagonian coast. Global aquaculture, which uses feeds manufactured from fish meal, increased by 50% between 1998 and 2004, and will likely continue to grow (*4*). Uruguay recently approved a Chilean-financed factory to process 200,000 tons of anchovy into fish meal (*5*). An increasing human population will create even greater demand for protein and nutrients derived from harvest of forage fish like anchovy.

The southwest Atlantic anchovy (*Engraulis anchoita*) is a crucial intermediate step in the flow of energy through the food web, dominating the level between tiny plankton and much of the wildlife of the Patagonian shelf (*1*). Commercially important fish and cephalopods, penguins, cormorants, terns, sea lions, and dolphins prey on the anchovy (*6*). Anchovy compose more than half the Magellanic penguin diet in the province of Chubut (*7*). The penguins also eat Argentine hake (*Merluccius hubbsi*), one of the commercially important fish species that prey on anchovy (*8*). Anchovy populations are naturally quite variable, and longer-lived predators are able to cope with this variability—as long as good years follow bad ones.

Food web interactions and trade-offs among competing fisheries require a multispecies management approach (*9*) if Argentina hopes to recover its hake fishery and simultaneously expand an anchovy fishery. Changes in anchovy populations can alter the abundance of both their predators and their prey. The effect of a decrease in the anchovy population could spread through the food web, changing the flow of energy and abundance of species not directly linked to the anchovy. These food web interactions are not yet quantitatively understood.

The spectacular wildlife of the Patagonian coast supports a thriving ecotourism industry. The Province of Chubut reported U.S.$165 million of direct revenue and U.S.$300 million of indirect revenue from tourism in 2005 (*10*), over half of which is associated with the biodiversity of the coast. If anchovy fishing reduced seabird numbers, especially of penguins, this revenue would be jeopardized.

Once a fishery is established, social pressures make it politically difficult to reduce fishing effort. The Argentine government declared a state of emergency for hake in 1999, when the hake fleet capacity exceeded the legal Total Allowable Catch (TAC) by a factor of three (*11*). The government faced stiff opposition to emergency fleet-specific bans to prevent further overfishing (*8*). Biologically rational decisions may not be politically possible once investment has occurred.

Argentine officials seek to provide employment and to generate revenue from an anchovy fishery. But before any further expansion and investment takes place, the costs to other fisheries, risks to wildlife and ecotourism, and food web interactions need to be determined. Costs and risks can then be weighed against the anticipated benefits under various management options. A conservative (precautionary) TAC, leaving a safety margin for natural fluctuations and unanticipated food web interactions, is needed to prevent overfishing and overinvestment. For adaptive management, data on ecosystem status, indicator species' populations over time, and food web interactions are needed to build quantitative understanding and to inform future management decisions.

[1]Department of Biology, University of Washington, Seattle, WA 98195, USA. [2]Fundación Patagonia Natural, Chubut 9120, Argentina.

*Correspondence. E-mail: skewes@u.washington.edu

References and Notes

1. A. Bakun, *Progr. Oceanogr.* **68**, 271 (2006).
2. Consejo Federal Pesquero, Resolución 6/03; www.cfp.gov.ar/resoluciones_2002_03.htm.
3. Secretaría de Agricultura, Ganadería, Pesca y Alimentos, "Desembarques de capturas marítimas totales"; www.sagpya.mecon.gov.ar (2006).
4. Food and Agriculture Organization (FAO), Fisheries Department, *The State of World Fisheries and Aquaculture (SOFIA) 2004* (FAO, Rome, 2004).
5. Presidencia de la República Oriental del Uruguay, "Se Adjudica Licitación Publica a la Firma Ibramar S.A. para Extracción y Procesamiento de la Especie Anchoita"; www.presidencia.gub.uy/resoluciones/2002072634.htm (2002).
6. M. Koen-Alonso, P. Yodzis, *Can. J. Fish. Aquat. Sci.* **62**, 1490 (2005).
7. E. Frere, P. Gandini, V. Lichtschein, *Ornitol. Neotrop.* **7**, 35 (1996).
8. S. N. Schonberger, J. J. Agar, *Argentina: Towards Rights-Based Fisheries Management* (Report no. 22816, World Bank, Washington, DC, 2001).
9. J. J. Magnuson *et al.*, *Dynamic Changes in Marine Ecosystems: Fishing, Food Webs, and Future Options* (National Academy Press, Washington, DC, 2006).
10. Gobierno del Chubut, "El turismo en Chubut generó ingresos directos por 500 millones de pesos"; www.chubut.gov.ar (2006).
11. FAO Fisheries Department, "Información sobre la ordenación pesquera de la República Argentina"; www.fao.org/fi/fcp/es/ARG/body.htm (2001).

12. We thank R. T. Paine, W. Conway, W. Wooster, R. Hilborn, and two anonymous reviewers for helpful comments. E. (Skewes) Skewgar was funded by an ARCS Foundation Fellowship, a U.S. Environmental Protection Agency STAR Fellowship, and an NSF Graduate Research Fellowship. This research was supported in part by the Wadsworth Endowed Chair in Conservation Science and the Wildlife Conservation Society. The views expressed are the authors' and not necessarily those of the sponsors.

10.1126/science.1135767

MOLECULAR BIOLOGY

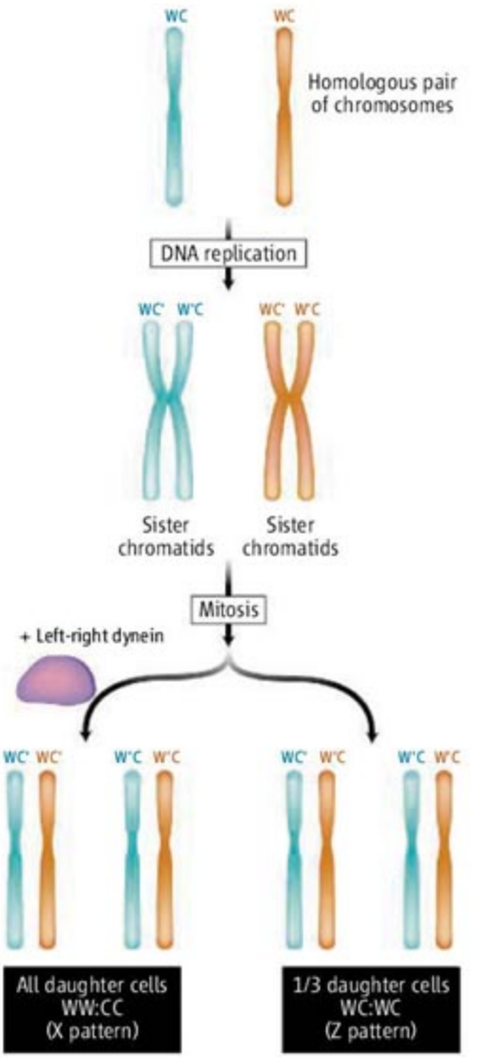
Do Watson and Crick Motor from X to Z?

A protein that determines left-right body asymmetry in the mouse is involved in the nonrandom segregation of duplicated chromosomes to daughter cells.

Carmen Sapienza

Eukaryotic diploid somatic cells reproduce by cell division (mitosis), in which each chromosome of a homologous pair (one from each parent) undergoes semiconservative DNA replication, producing a copy of each homolog (see the figure). After replication and chromosome condensation, microtubules belonging to a structure called the mitotic spindle attach to opposite sides of each replicated homolog and pull one of the two copies (sister chromatids) to opposite poles. Barring uncorrected replication errors, the semiconservative nature of DNA replication ensures that each sister chromatid is identical and that each daughter cell will be genetically identical to the parent cell. Given this identity, most biologists believe that which chromatid segregates to which daughter cell is immaterial. However, having two copies of each homolog does bring up the potentially vexing issue of choice. Are both copies equally good? How is that decided? If they are not equally good, then what happens? On page 100 of this issue, Armakolas and Klar (*1*) start to address these questions, although which question is actually addressed is likely to be the subject of debate—how chromatids are distinguished versus how they are segregated.

One can imagine situations in which the choice of which chromatid to segregate to which daughter cell might make a difference. Cairns (*2*) proposed that it would be advantageous to segregate the "oldest" DNA strands—that is, the original DNA, as opposed to new DNA that is synthesized during replication—to the stem cell daughter in any division that produced both a stem cell and a differentiated cell. Keeping the oldest strands in the stem cell would reduce the possibility that replication errors might affect the stem cell population and might reduce the risk of cancer. Another opportunity to put strand identity to good use has been envisioned by Klar (*3*), who argued that strand-specific imprinting and patterned segregation of DNA strands during mitosis could be the basis for forming the left-right body axis during development. In this model, nonrandom chromatid segregation arises when chromatids containing the old "Watson" (W) DNA strands segregate into one daughter cell while chromatids containing the old "Crick" (C) DNA strands segregate into the other daughter cell—in other words, a WW:CC segregation pattern. In fact, this specific proposal by Klar, in combination with the results of earlier work (*4*), has led to the present report that identifies a factor involved in biased segregation of chromatids during mitosis.



Motors and daughters. During cell division, Armakolas and Klar (*1*) propose that replicated sister chromatids can segregate into daughter cells nonrandomly (all daughters having a WW:CC, or X segregation pattern) rather than randomly (daughters having a mixture of WW:CC and WC:WC, or X and Z segregation patterns). A motor protein (left-right dynein) may influence sister chromatid segregation. W and C designate older DNA strands, whereas W' and C' represent newly synthesized DNA strands.

Armakolas and Klar have used an established mouse cell culture system (*5*) in which it is possible to distinguish the segregation of sister chromatids of mouse chromosome 7. In this experimental system, a mitotic recombination event is induced that reconstitutes a drug resistance gene (*Hprt*) on only one of the two chromatids involved in the recombination event. Thus, drug selection produces cells that carry the *Hprt*-bearing recombinant chromatid from one homolog in all cases. To test whether segregation of chromatids is random or not, one need only determine which chromatid of the homologous chromosome 7 segregates to the drug-resistant cell—the nonrecombinant chromatid (called an X segregation pattern) or the recombinant chromatid (a Z segregation pattern). These correspond to the WW:CC segregation pattern and the WC:WC pattern, respectively, in the model proposed by Armakolas and Klar.

In this experimental system, the prevailing view on the segregation of chromatids during mitosis is that the X mode (WW:CC) is predominant and results from physical constraints imposed on the mitotic chiasma (the physical point of crossover between two chromatids that facilitates exchange of pieces of chromatid) and by sister chromatid cohesion (*5*, *6*). That being said, "predominant" does not mean "exclusive," and herein lies the intellectual root of Armakolas and Klar's experiment. Liu *et al.* (*5*) and Armakolas and Klar (*4*) reported exclusive (100%) cosegregation of the reconstituted drug-resistance gene with the nonrecombinant chromatid from the homolog (X segregation) in a mouse embryonic stem cell system. Armakolas and Klar also described exclusive X and Z segregation in an endoderm and neuroectoderm cell line, respectively (*4*). They proposed that the exclusive segregation modes result from biased (nonrandom) segregation of DNA strands from each homolog to each daughter cell and that these patterns are cell-type specific (*4*). Although well-reasoned objections have been raised to this explanation (*7*) (the present

The author is at the Fels Institute for Cancer Research and Department of Pathology, Temple University Medical School, 3307 North Broad Street, Philadelphia, PA 19140, USA. E-mail: sapienza@temple.edu

results do not shed any direct light on this controversy), Armakolas and Klar carried their supposition one step further: If the factors that influence segregation of DNA strands are the same factors that influence left-right body axis formation, then how might a gene product that influences body axis formation influence the segregation of chromatids? They focused on the gene encoding the left-right dynein motor protein (LRD). Mutations in the mouse gene (*Dnahc11*) and the human homolog (*DNAH11*) encoding this motor protein cause left-right axis randomization of some internal organs.

When Armakolas and Klar used the same *Hprt*-recombination experimental system, and reduced expression of the left-right dynein motor by RNA interference, chromatid segregation became nearly "random" in those cell lines in which it had been exclusively the X or Z type. All three cell lines reverted to predominantly the X segregation pattern regardless of whether they were 100% X (embryonic stem cells and endoderm cells) or 100% Z (neuroectoderm cells) in the first place. This segregation ratio (2:1 to 3:1 X:Z segregants) in the absence of LRD is the same as that observed in cell lines that do not normally express LRD (pancreatic cells, mesoderm cells, and cardiomyocytes) and approximately the same as that reported for other embryonic stem cell lines (*5*) for which the status of LRD expression is unknown. Interestingly, this X:Z segregant ratio is also observed in the fruit fly *Drosophila melanogaster* (*6*).

One explanation for this ratio is that segregation of *Hprt*-recombinant chromatids presents a topological problem with a single solution (X segregation). Apparent instances of Z segregation are thought to arise as a result either of recombination that normally occurs before DNA replication or of recombination between homologs and between sister chromatids followed by X segregation. The most perplexing observation is not what happens in the absence of LRD but why the presence of LRD leads to exclusive X or Z segregation. What could LRD be doing?

There are at least two possibilities. In neuroectoderm cells, LRD could eliminate *Hprt*-recombination in the G2 phase of the cell division cycle, which is just before the onset of mitosis (but after DNA replication has occurred). On the other hand, in embryonic stem cells and endoderm cells, LRD could eliminate *Hprt*-recombination during the G1 phase, which is before DNA replication begins. This explanation requires only that LRD have a strong negative effect on recombination but does so at different times during the cell cycle in different types of cells.

The second possibility is that LRD directly affects the orientation of the joined homologs on the spindle, placing the *Hprt*-recombinant chromatids on opposite sides of the metaphase plate (the region of the mitotic spindle where replicated chromosomes are positioned before separation of chromatids into daughter cells) in embryonic stem cells and endoderm cells (X segregation) or on the same side of the metaphase plate (Z segregation) in neuroectoderm cells. It is unclear how LRD might play such a chromosome-orientation role and how the decision on which orientation to take could be based on strand identity. Nevertheless, it is suspicious that a dynein motor protein—a family whose members are involved in chromosome movement—affects chromatid segregation.

Regardless of how this phenomenon is ultimately explained, Armakolas and Klar are to be commended for testing an unorthodox hypothesis by an experiment that was not an obvious approach. Major scientific discoveries are rarely accompanied by investigators shouting "eureka" but are often accompanied by investigators mumbling "that's strange." At first sight, I confess I thought it a strange result.

References

1. A. Armakolas, A. J. S. Klar, *Science* **315**, 100 (2007).
2. J. Cairns, *Nature* **255**, 197 (1975).
3. A. J. S. Klar, *Trends Genet.* **10**, 392 (1994).
4. A. Armakolas, A. J. S. Klar, *Science* **311**, 1146 (2006).
5. P. Liu, N. A. Jenkins, N. G. Copeland, *Nat. Genet.* **30**, 66 (2002).
6. K. J. Beumer, S. Pimpinelli, K. G. Golic, *Genetics* **150**, 173 (1998).
7. J. E. Haber, *Science* **313**, 1045b (2006).

10.1126/science.1137587

PHYSICS

Negative Refractive Index at Optical Wavelengths

Costas M. Soukoulis, Stefan Linden, Martin Wegener

Metamaterials are designed to have structures that provide optical properties not found in nature. If their capacity can be extended, new kinds of devices for imaging and control of light will be possible.

Although discovered only 6 years ago, negative refractive index materials (NIMs) have been the target of intense study, drawing researchers from physics, engineering, materials science, optics, and chemistry. These artificial "metamaterials" are fascinating because they allow the design of substances with optical properties that simply do not occur in nature (*1*–*4*). Such materials make possible a wide range of new applications as varied as cloaking devices and ultrahigh-resolution imaging systems. The variety of possible applications would be even greater if such materials could be engineered to work at optical wavelengths.

For the ultimate control of light, one needs a handle on both the electric and the magnetic components of the electromagnetic (EM) light wave. To achieve this control, normally one would think about modifying the microscopic electric and magnetic fields in a material. However, in most cases it is easier to average over the atomic scale and consider the material to be a homogeneous medium characterized by the electric permittivity ε and the magnetic permeability μ. These two quantities describe the EM response of a given material. More specifically, Veselago showed nearly 40 years ago (*5*) that the combination $\varepsilon < 0$ and $\mu < 0$ leads to a negative refractive index, $n < 0$. This means that the phase velocity of light is negative; in other words, light waves now have a "reverse gear."

Veselago's idea remained obscure because no such natural materials were known to exist at any frequency. Although electric resonances with $\varepsilon < 0$ do occur up to the visible and beyond, magnetic resonances typically die out at microwave frequencies. Moreover, the electric and magnetic resonances would need to overlap in frequency, which seemed improbable. However, by making use of artificially structured metamaterials, in which inclusions smaller than a

C. M. Soukoulis is at the Ames Laboratory and Department of Physics and Astronomy, Iowa State University, Ames, IA 50011, USA, and at the Institute of Electronic Structure and Laser, Foundation for Research and Technology–Hellas, 71110 Heraklion, and the University of Crete, 71409 Heraklion, Crete, Greece. S. Linden and M. Wegener are at the Center for Functional Nanostructures, Universität Karlsruhe and Forschungszentrum Karlsruhe, D-76128 Karlsruhe, Germany. E-mail: soukoulis@ameslab.gov

wavelength replace the atoms and molecules of a conventional material, scientists can circumvent this limitation. Metamaterials can be designed to exhibit both electric and magnetic resonances that can be separately tuned to occur in spectra from the low radio-frequency to the visible.

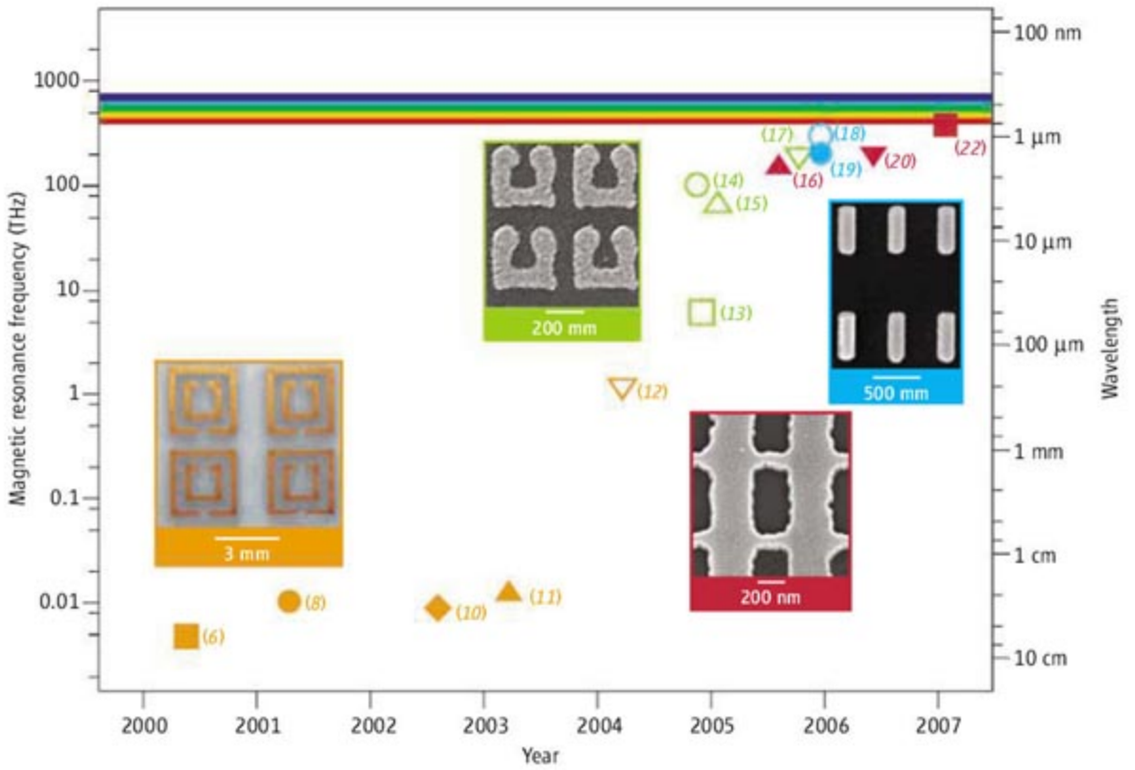
Since the first demonstration (*6*) of an artificial NIM in 2000, metamaterials have exhibited a broad range of properties and potential applications: nearly zero reflectance; nanometer-scale light sources and focusing; miniaturization of devices, such as antennas and waveguides; and novel devices for medical imaging, especially magnetic resonance imaging. For example, metamaterials may lead to the development of a flat superlens (*7*) that operates in the visible spectrum, which would offer superior resolution over conventional technology and provide image resolutions much smaller than one wavelength of light.

Subsequent theory and experiment (*8*–*22*) confirmed the reality of negative refraction. The development of NIMs at microwave frequencies (*6*, *8*–*11*) has progressed to the point where scientists and engineers are now vigorously pursuing microwave applications. In contrast, research on NIMs that operate at higher frequencies (*12*–*22*) is at an early stage, with issues of material fabrication and characterization still being sorted out.

The figure gives a detailed history of the development of the magnetic resonance frequency and/or the frequency of negative n as a function of time. In the early years of the field (2000 to 2003), the design of choice to obtain $\mu < 0$ was an artificial structure proposed by Pendry, the so-called split-ring resonator (SRR). This structure exhibits a band of negative μ values even though it is made of nonmagnetic materials. A double SRR is shown at the lower left of the figure. A negative μ at 10 GHz requires SRR dimensions on the order of 1 mm. To obtain negative ε, one needs to arrange long and thin wires in a simple cubic lattice, so as to mimic the response of a metal to electromagnetic waves—that is, below a frequency called the plasma frequency, ε is negative. Negative ε at gigahertz frequencies might be obtained with wires a few tens of micrometers in diameter and spaced several millimeters apart. By using an array of SRRs and thin wires in alternating layers, several groups (*6*, *8*–*11*) showed negative n at gigahertz frequencies.

As can be seen from the figure, a negative μ at terahertz and infrared frequencies was achieved in 2004. The idea underlying that work was that the magnetic resonance frequency of the SRR is inversely proportional to its size. Thus, the concepts from the microwave regime could simply be scaled down to shorter wavelengths. For ease of fabrication, a transition from double SRR to single SRR took place (see the figure). Indeed, this approach works up to about 200 THz. Unfortunately, it was found that this scaling breaks down for yet higher frequencies for the single SRR. The reason is that the metal of which the SRR is composed starts to strongly deviate from an ideal conductor.

Although these developments have been important proofs of principle, progress was hindered by several experimental details. For example, the combination of these SRRs with metal wires to form a three-dimensional structure is very challenging on the nanometer scale. Thus, there was a hunt for alternative designs that are more suitable for the terahertz or even for the visible regime. The key idea to make this possible was independently realized and published by three different groups in 2005 (*16*, *18*, *19*). These designs all show that pairs of metal wires or metal plates, separated by a dielectric spacer, can provide the magnetic resonance. The magnetic resonance originated from the antiparallel current in the wire pair with an opposite sign charge accumulating at the corresponding ends. This resonance provides $\mu < 0$. In addition, an electric resonance with $\varepsilon < 0$ results for excitation of a parallel current oscillation. In the transmission measurements, the EM waves were incident normal to the sample surface. This setup is much simpler than that for conventional SRRs and wires, where the incident EM waves must propagate parallel to the sample surface.



Advances in metamaterials. The solid symbols denote $n < 0$; the open symbols denote $\mu < 0$. Orange: data from structures based on the double split-ring resonator (SRR); green: data from U-shaped SRRs; blue: data from pairs of metallic nanorods; red: data from the "fishnet" structure. The four insets give pictures of fabricated structures in different frequency regions.

Overlap (*18*, *19*) of the regions where ε and μ are both negative with only wire pairs is difficult, so new designs were needed. One way is to introduce extra continuous wires next to the pairs, or to change the shape of the wires. The best design that has been used in 2005 and 2006 is the so-called "double-fishnet" structure, which consists of a pair of metal fishnets separated by a dielectric spacer. This design is shown in the lower right of the figure. Although the choice of the metal constituting the structure is not critical in the microwave regime, it is crucial in the optical and the visible regime because the metamaterial losses are dominated by metal losses. Silver exhibits the lowest losses at optical frequencies, and indeed, going from gold (*16*, *18*, *20*) to silver drastically reduced the losses at similar fre-

quencies (*21*). A suitable measure for the losses is the figure of merit (FOM), defined as the negative ratio of the real to the imaginary part of n. Dolling *et al.* (*21*) obtained FOM = 3 at a wavelength of 1400 nm, which compares to FOM < 1 for other groups (*16*, *19*, *20*). Furthermore, the use of silver has enabled the first negative-index metamaterials at the red end of the visible spectrum (*22*) (wavelength 780 nm). Another group has also reported a negative n (*23*, *24*), but this has been questioned recently (*25*).

Only 6 years after their first demonstration, negative-index metamaterials have been brought from microwave frequencies toward the visible regime. However, for applications to come within reach, several goals need to be achieved: reduction of losses (by using crystalline metals and/or by introducing optically amplifying materials), three-dimensional rather than planar structures, isotropic designs, and ways of mass production of large-area structures. With emerging techniques such as microcontact printing, nanoembossing, holographic lithography, and quantum tailoring of large molecules, it seems likely that these technical challenges can be successfully met. The spirit of metamaterials is to design materials with new and unusual optical properties. In that enterprise, only our imagination and creativity set the limits.

References and Notes

1. D. R. Smith, J. B. Pendry, M. C. K. Wiltshire, *Science* **305**, 788 (2004).
2. D. R. Smith, J. B. Pendry, *Phys. Today* (June 2004), p. 37.
3. C. M. Soukoulis, *Opt. Phot. News* (June 2006), p. 16.
4. C. M. Soukoulis, M. Kafesaki, E. N. Economou, *Adv. Mater.* **18**, 1941 (2006).
5. V. G. Veselago, *Sov. Phys. Uspekhi* **10**, 509 (1968).
6. D. R. Smith *et al.*, *Phys. Rev. Lett.* **84**, 4184 (2000).
7. J. B. Pendry, *Phys. Rev. Lett.* **85**, 3966 (2000).
8. R. A. Shelby, D. R. Smith, S. Schultz, *Science* **292**, 77 (2001).
9. C. G. Parazzoli *et al.*, *Phys. Rev. Lett.* **90**, 107401 (2003).
10. M. Bayindir *et al.*, *Appl. Phys. Lett.* **81**, 120 (2002).
11. R. B. Greegor *et al.*, *Appl. Phys. Lett.* **82**, 2356 (2003).
12. T. J. Yen *et al.*, *Science* **303**, 1494 (2004).
13. N. Katsarakis *et al.*, *Opt. Lett.* **30**, 1348 (2005).
14. S. Linden *et al.*, *Science* **306**, 1351 (2004).
15. S. Zhang *et al.*, *Phys. Rev. Lett.* **94**, 037402 (2005).
16. S. Zhang *et al.*, *Phys. Rev. Lett.* **95**, 137404 (2005).
17. C. Enkrich *et al.*, *Phys. Rev. Lett.* **95**, 203901 (2005).
18. G. Dolling *et al.*, *Opt. Lett.* **30**, 3198 (2005).
19. V. M. Shalaev *et al.*, *Opt. Lett.* **30**, 3356 (2005).
20. G. Dolling, C. Enkrich, M. Wegener, C. M. Soukoulis, S. Linden, *Science* **312**, 892 (2006).
21. G. Dolling *et al.*, *Opt. Lett.* **31**, 1800 (2006).
22. G. Dolling *et al.*, *Opt. Lett.* **32**, 53 (2007).
23. A. N. Grigorenko *et al.*, *Nature* **438**, 335 (2005).
24. A. N. Grigorenko, *Opt. Lett.* **31**, 2483 (2006).
25. A. V. Kildishev *et al.*, http://arxiv.org/abs/physics/0609234 (2006).

26. We thank Th. Koschny and J. Zhou for preparing the figure. Supported by Ames Laboratory, operated by Iowa State University under contract W-7405-Eng-82 (C.M.S.), and by Helmholtz-Hochschul-Nachwuchsgruppe grant VH-NG-232 (S.L.).

10.1126/science.1136481

ECOLOGY

The Heartbreak of Adapting to Global Warming

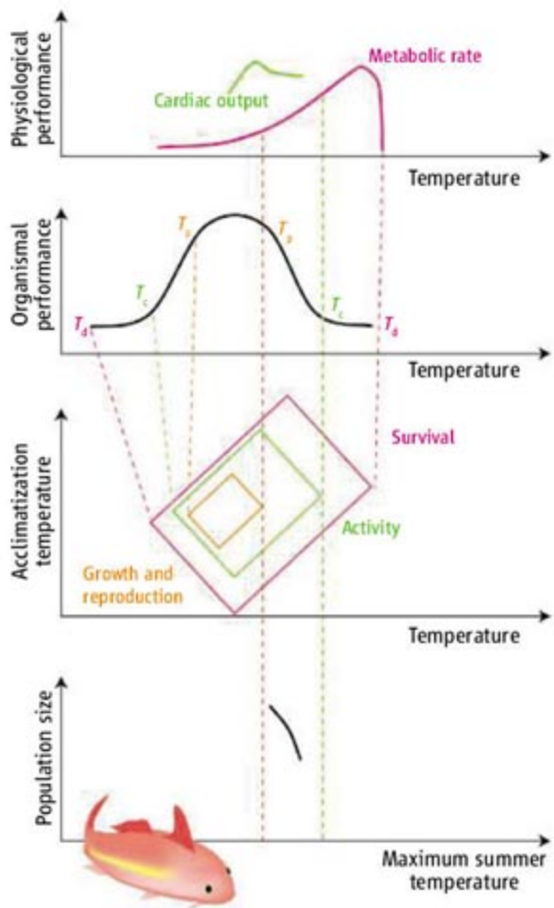
Laboratory studies of basic physiological constraints on the cardiorespiratory system can be used to predict the impact of global warming on fish.

Tobias Wang and Johannes Overgaard

Climatic changes have been linked to altered geographical distributions of many organisms, including marine fish (*1*, *2*). Yet, it remains difficult to distinguish direct causal relations between environmental temperature and species distribution patterns (*3*) from indirect effects through interactions with prey, predators, pathogens, or competitors (*4*). An ambitious goal of integrative biology is to understand how temperature affects physiological mechanisms at all levels of biological organization. This could allow predictions of how global warming affects animal performance and population dynamics. Animal physiologists commonly rely on laboratory studies to predict temperature tolerance of animals, but whole-animal performance in natural settings is rarely investigated. On page 95 of this issue, Pörtner and Kunst (*5*) provide compelling evidence that thermal constraints on oxygen transport are causing the population of a marine fish, the viviparous eelpout (*Zoarces viviparus*), to decline in the Wadden Sea.

T. Wang is in the Department of Zoophysiology, Aarhus University, 8000 Aarhus, Denmark. E-mail: tobias.wang@biology.au.dk]. Overgaard is at the National Environmental Research Institute, Department of Terrestrial Ecology, 8600, Silkeborg, Denmark.

Over the past decade, Pörtner and co-workers have studied various aspects of oxygen transport and metabolism in numerous animal species, including the viviparous eelpout (*6*). They have identified the pejus temperature (pejus means "turning worse"), beyond which the ability of animals to increase aerobic metabolism is reduced. This reduction is evident from the decline in aerobic scope, which is defined as the proportional difference between resting and maximal rates of oxygen consumption. The temperature range between the lower and higher



Thermal limits. Beyond the "pejus" temperature (T_p), the cardiorespiratory system of the fish can no longer ensure sufficient aerobic scope to sustain reproduction and growth; eventually, activity (T_c) and survival (T_d) are also compromised. These thermal limits are plastic and amenable to the thermal history (acclimatization) of the animals (**top three panels**). Pörtner and Kunst show that summer temperatures above the pejus cause the population of the European eelpout to decline, indicating that global warming may take effect well before the lethal thermal limits (T_d) are reached (**bottom panel**).

pejus temperatures is much narrower than that between the critical temperatures (T_c), beyond which the animal only survives for short periods (see the figure).

As in other animals, continued cardiac function is essential in fish, but coronary circulation is normally sparse. Thus oxygen to the fish heart is primarily provided by the venous blood returning from the body (*7*). The oxygen concentration of venous blood declines if cardiac output does not increase in proportion to the rise in metabolism that occurs with elevated temperature (*8*). These problems are exacerbated by the fact that the concentration of physically dissolved oxygen in the water declines progressively with increased temperature. As a result, the heart is likely to limit the aerobic scope, rendering the fish more vulnerable to predators and less effective as a forager.

The novel discovery of Pörtner and Kunst is their observation of a strong negative correlation between estimated population sizes and summer temperatures over the past ~50 years. On a shorter time scale, the authors also found that warm summers strongly reduced population size the following year. It remains difficult to establish increased temperature as the mechanistic cause for the population decline, but the correlation to the pejus and critical threshold temperatures derived from laboratory data is persuasive.

The temperatures causing population declines are considerably lower than the critical temperatures. The population appears to decline before temperature threatens survival of the individual. Thus, lowered scope for growth and reproduction, rather than heat-induced death per se, appears to cause the population decline.

A potential limitation of the study by Pörtner and Kunst is the difficulty of assessing the role of acclimatization. The temperatures to which an animal has previously been exposed can improve its ability to survive heat and cold, and can affect the thermal thresholds at both low and high temperature (*9*). The tight correlation between summer temperatures and population size observed by the authors may, nevertheless, indicate that such thermal adaptation is exhausted for eelpout in their most southern distribution range. Indeed, a lack of an acclimatory response in marine animals has previously been correlated with the inability to handle thermal shifts (*2*, *10*).

Population dynamics are complex and depend on many biological and physical parameters, but as shown by Pörtner and Kunst, a thorough understanding of physiological limitations may provide the necessary insight to determine how global warming affects animal performance (*11*). The association between thermal tolerance of the oxygen transport system and population declines shows that old-fashioned physiology can be essential for understanding how temperature determines the geographical distributions of animals.

References

1. G. R. Walther *et al.*, *Nature* **416**, 389 (2002).
2. A. L. Perry, P. J. Low, J. R. Ellis, J. D. Reynolds, *Science* **308**, 1912 (2005).
3. M. N. Jensen, *Science* **299**, 38 (2003).
4. A. J. Davis *et al.*, *Nature* **391**, 783 (1998).
5. H. O. Pörtner, R. Kunst, *Science* **315**, 95 (2007).
6. M. Frederich, H. O. Portner, *Am. J. Physiol.* **283**, R1531 (2000).
7. A. P. Farrell, in *The Vertebrate Gas Transport Cascade: Adaptations to Environment and Mode of Life*, E. Bicudo, Ed. (CRC Press, Boca Raton, FL, 1993), pp. 208–214.
8. A. G. Heath, G.M. Hughes, *J. Exp. Biol.* **59**, 323 (1973).
9. A. R. Cossins, K. Bowler, in *Temperature Biology of Animals* (Chapman & Hall, London, 1987), pp. 210–220.
10. J. H. Stillman, *Science* **301**, 65 (2003).
11. B. Helmuth, J. G. Kingsolver, E. Carrington, *Annu. Rev. Physiol.* **67**, 177 (2005).

10.1126/science.1137359

ATMOSPHERE

Aerosols Before Pollution

No unpolluted regions remain in today's atmosphere. How can we estimate the aerosol content of the atmosphere before there was human activity?

Meinrat O. Andreae

Atmospheric aerosols play a large role in human-induced climate change because of their effects on solar radiation transfer and cloud processes. To assess the impact of human perturbations on the atmosphere's aerosol content, we need to know the prehuman aerosol burden. This is especially important for understanding the cloud-mediated effects of aerosols on climate, because cloud properties respond to aerosols in a nonlinear way and are most sensitive to the addition of particles when the background concentration is very low (*1*). Because cloud droplets can nucleate only on particles above a certain size (typically about 60 to 90 nm), this subset of the aerosol population—called cloud condensation nuclei (CCN)—is of particular importance. In the following, I try to provide a rough estimate of what CCN concentrations might have been in the prehuman atmosphere.

The author is at the Max Planck Institute for Chemistry, 55020 Mainz, Germany. E-mail: andreae@mpch-mainz.mpg.de

Information about atmospheric aerosol contents in the absence of human activity is very difficult to obtain. Human activities are causing the emission of huge amounts of aerosol particles and their gaseous precursors. Aerosol particles have typical atmospheric lifetimes of 3 to 10 days; on average, after three such lifetimes, about 5% of the initial burden remains in the atmosphere. Given that air masses can easily travel several thousand kilometers in 15 days, there are really no places where we can expect to find truly pristine conditions, especially in the Northern Hemisphere.

Aerosol concentrations approaching pristine conditions are mostly found over the oceans, especially in the Southern Hemisphere, where large expanses of open ocean and a low density of population and industry contribute to keeping the human impact at a minimum. The natural aerosol over these remote ocean regions consists mainly of a mixture of sea salt particles, organics, and sulfates from the oxidation of biogenic dimethylsulfide; some mineral dust and smoke from wildfires may also be present (see the figure). In biologically productive ocean regions, typical concentrations of CCN are in the low hundreds per cm^3. Much lower concentrations of a few tens of CCN per cm^3 are found over the mid-latitude oceans in wintertime, when biological and photochemical activity are low.

The determination of pristine CCN concentrations over continental regions presents a much more difficult problem. Measurements at sites away from obvious sources of pollution are very few, and even among these data, it is usually difficult to assess how much of the observed aerosol results from pollution. Aerosol compositions at remote sites in the Northern Hemisphere suggest that the continental "background" aerosol nowadays consists mostly of pollution aerosols at varying levels of dilution: The concentration of black carbon, a unique indica-

tor of combustion and pollution, is strongly correlated to that of the dominant sulfate and organic aerosols (*2*). Even in the Southern Hemisphere, pollution aerosols, especially from biomass burning, dominate in most continental areas, with CCN concentrations typically in the upper hundreds to thousands per cm^3.

Over remote continental regions, the cleanest conditions prevail when unpolluted air masses of marine origin flow over nearly uninhabited lands. For example, measurements have been made in the center of the Amazon Basin during the rainy season, when clean air masses from the Atlantic Ocean are transported for several days over the Amazon forest. CCN concentrations were in the low hundreds per cm^3, more or less identical to the concentrations over the tropical oceans (*3*). Similar concentrations have been reported from other remote continental sites, such as southeast Australia, the western United States and Alaska, and northern Finland (*4*–*7*). Clearly, all these measurements represent upper limits to the natural CCN populations, because even these locations are influenced to varying degrees by the long-range transport of pollution.

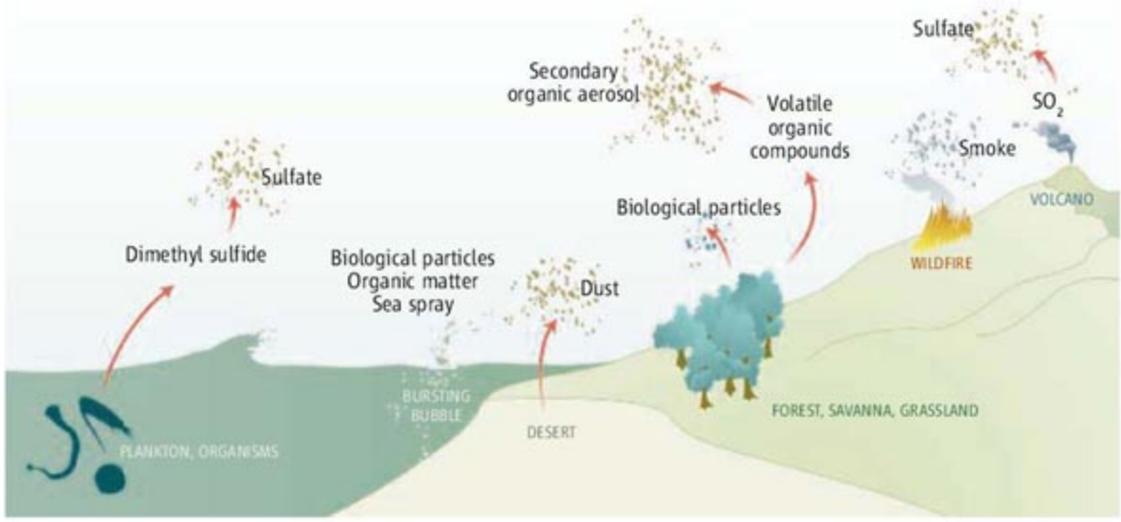
An alternative way of assessing the pristine continental CCN background is by estimating the number of new particles in the CCN size range produced from biogenic precursors at remote sites. During summer, bursts of particle production occur in such places about twice a week, but this mechanism cannot sustain a substantial CCN population on a continuous basis. To get a more representative perspective on aerosol particle formation, Tunved *et al.* (*8*) determined the increase in the number of particles as air masses traveled from the Atlantic over land to research sites in Finland. Particle numbers increased with travel time and the rate of terpene emission from plants. At typical terpene emission rates, total particle concentrations of ~1000 to 2000 per cm^3 were reached, of which ~100 to 300 were larger than 90 nm and therefore potential CCN.

These data are from a region where nucleation is favored because of trace amounts of anthropogenic SO_2, and they only apply to the spring and summer seasons. Thus, they probably still represent an upper limit to natural CCN production at mid-latitudes. Overall, natural production of CCN-active particles over biologically active regions on the continents probably cannot account for more than 100 to 300 per cm^3, not much greater than the levels found over the oceans. During the cold seasons, much lower particle production must be expected.

In recent years, modelers have tried to reproduce pristine aerosol conditions by running their global chemistry/transport/climate models with industrial or anthropogenic sources turned off (*9*). Unfortunately, the production rates and mechanisms for primary biogenic aerosols (plant particles, spores, microbes, etc.) and secondary organic aerosols (from natural hydrocarbons) are still very poorly understood. These two components may be responsible for a large fraction of the natural continental aerosol, and current model results can therefore only be considered rough estimates of preindustrial aerosol abundance over the continents. This applies especially to number concentrations and size distributions, which are our primary concern here.

I am not aware of any modeling studies that have attempted to look at the atmosphere before the advent of humans. Instead, the models use as a reference state either the preindustrial period or the present-day atmosphere with anthropogenic sources turned off. All models agree that anthropogenic emissions have caused large enhancement of aerosol loads even over remote parts of the continents, with typical enhancements by 50 to 300% over remote regions of Asia, North America, and South America. From these studies, we can estimate preindustrial CCN concentrations over the continents of 50 to 200 per cm^3, similar to the values over the remote oceans in the same models. Higher aerosol concentrations are predicted over the tropical continents, because of biomass burning by preindustrial human populations.

Thus, prehuman aerosol levels may have been very similar over continents and oceans, ranging from a few tens per cm^3 in biogenically inactive regions or seasons to a few hundreds per cm^3 under biologically active conditions. This conclusion renders invalid the conventional classification of air masses into maritime and continental according to their aerosol content. It also implies that, before the onset of human-induced pollution, cloud microphysical properties over the continents resembled those over the oceans, whereas nowadays, cloud processes over most of the continents are shaped by the effects of human perturbation.



Sources of aerosol particles to the natural atmosphere. Primary particles—such as sea spray, soil dust, smoke from wildfires, and biological particles including pollen, microbes, and plant debris—are emitted directly into the atmosphere. Secondary particles are formed in the atmosphere from gaseous precursors; for example, sulfates form from biogenic dimethyl sulfide and volcanic sulfur dioxide (SO_2), and secondary organic aerosol from biogenic volatile organic compounds.

References

1. U. Lohmann, J. Feichter, *Atmos. Chem. Phys.* **5**, 715 (2005).
2. J. P. Putaud *et al.*, *Atmos. Environ.* **38**, 2579 (2004).
3. G. C. Roberts, M. O. Andreae, J. Zhou, P. Artaxo, *Geophys. Res. Lett.* **28**, 2807 (2001).
4. S. Twomey, K. A. Davidson, K. J. Seton, *J. Atmos. Sci.* **35**, 650 (1978).
5. D. J. Delene, T. Deshler, *J. Geophys. Res.* **106**, 12579 (2001).
6. W. A. Hoppel, J. E. Dinger, R. E. Ruskin, *J. Atmos. Sci.* **30**, 1410 (1973).
7. M. Komppula, H. Lihavainen, V. M. Kerminen, M. Kulmala, Y. Viisanen, *J. Geophys. Res.* **110**, D06204 (2005).
8. P. Tunved *et al.*, *Science* **312**, 261 (2006).
9. P. Stier, J. Feichter, E. Roeckner, S. Kloster, M. Esch, *Atmos. Chem. Phys.* **6**, 3059 (2006).

10.1126/science.1136529

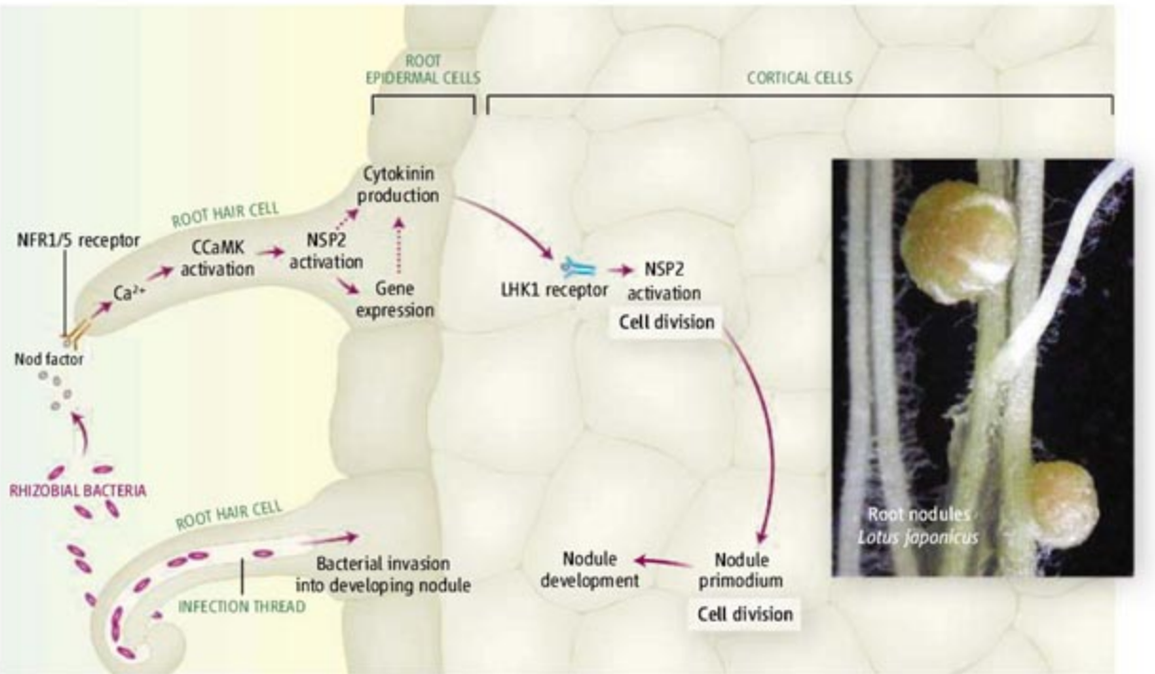
PLANT SCIENCE

Nodules and Hormones

A hormone-signaling pathway is crucial to the ability of certain plants to form nodules when stimulated by nitrogen-fixing bacteria.

Giles E. D. Oldroyd

Since the advent of the "green revolution," agriculture has been highly reliant on industrial nitrogen fertilizers to maximize crop production by enhancing the amount of nitrogen in soils. This comes at a heavy price. The industrial fixation of nitrogen accounts for ~50% of fossil fuel usage in agriculture. The leaching of fertilizers into aquatic systems accounts for increasing plant and algal blooms, now a global problem. The rising cost of nitrogen fertilizers—driven by the rising cost of fossil fuels—and the need for improved sustainability are making alternatives to nitrogen fertilizers ever more attractive. Two reports in this issue, one by Murray *et al.* on page 101 (*1*) and the other by Tirichine *et al.* on page 104 (*2*), provide insights into a natural replacement for industrial nitrogen fertilizers.



Nodule organogenesis. Nod factor, which is released by rhizobial bacteria, is perceived by the plant root epidermis. This triggers a calcium-dependent signaling pathway and the production of the hormone cytokinin. Localized increase in cytokinin may activate cortical cell division, leading to formation of the nodule primordium. Rhizobial bacteria infect the nodule through infection threads that are initiated in root hair cells and invade the developing nodule.

A number of plant species have developed an intimate association with nitrogen-fixing rhizobial bacteria that provide plants with a reliable nitrogen source. Central to this plant-bacterial interaction is the formation of unique organs on plant roots called nodules, which accommodate the bacteria and provide a suitable environment for nitrogenase, the bacterial enzyme responsible for nitrogen fixation. Nodules are plant-derived structures that form following recognition of rhizobial bacteria by the plant. Murray *et al.* and Tirichine *et al.* reveal that activation of a plant hormone signaling pathway in the legume *Lotus japonicus*, most likely by rhizobial bacteria, is sufficient to activate nodule formation. This discovery could pave the way to transferring this symbiotic process into other plant species.

It has long been known that the plant hormone cytokinin is important for nodule development. One of the earliest indicators of this was the observation that transfer of cytokinin production allows normally nonsymbiotic bacteria to activate nodule formation in alfalfa (*3*). This suggested that increased amounts of cytokinin could induce nodule formation. It was also reported that expression of a gene engineered to be responsive to cytokinin is activated in nodule primordia (*4*), and suppression of a cytokinin receptor expression by RNA interference, reduces nodulation in the legume *Medicago truncatula* (*5*). Curiously, treatment of roots with exogenous cytokinin activates expression of the nodulation gene *ENOD40* (*6*, *7*), but fails to induce nodule formation. Thus, cytokinin is involved in nodulation, but whether this plant hormone alone is sufficient to activate nodule formation has not been clear.

Murray *et al.* and Tirichine *et al.* resolve this question by demonstrating that activation of a cytokinin receptor, lotus histidine kinase 1 (LHK1), is both necessary and sufficient for nodule organogenesis (see the figure). Both studies use a genetic approach to provide complementary evidence for the central role of this cell surface receptor in nodule development. Murray *et al.* demonstrate that loss-of-function mutations in the *LHK1* gene of *L. japonicus* abolish nodule primordium formation, but do not affect bacterial invasion of the root. In these mutants, infection threads in root cells, which enable bacterial invasion, become highly elaborated and lose normal targeted growth toward the developing nodule primordia. Conversely, Tirichine *et al.* show that a gain-of-function mutation in *LHK1* leads to spontaneous formation of nodules in the absence of rhizobial bacteria. These mutant plants show additional phenotypes indicative of a constitutive cytokinin response and also show hypersensitivity to exogenous cytokinin.

Both studies reveal that LHK1 functions as a cytokinin receptor by expressing this protein in a heterologous system (bacteria or yeast). In these tests, the gain-of-function *LHK1* mutant shows constitutive cytokinin-independent activity. Activation of LHK1 thus appears critical for nodule organogenesis, but bacterial entry through infection thread formation appears to be independent of cytokinin signaling.

During nodule formation, there must be coordinated development between the root epidermis and the root cortex. Cortical cell divisions that define the nodule meristem only occur in regions immediately below epidermal cells where infection threads are initiat-

The author is in the Department of Disease and Stress Biology, John Innes Centre, Norwich Research Park, Norwich NR4 7UH, UK. E-mail: giles.oldroyd@bbsrc.ac.uk

PHOTO CREDIT: ANNE HECKMANN/JOHN INNES CENTRE

ing. Most epidermal responses are activated by the secreted rhizobial signaling molecule Nod factor. This molecule activates a signaling pathway in root epidermal cells that leads to oscillations in the intracellular concentration of calcium. This calcium signal is perceived by a calcium- and calmodulin-dependent protein kinase (CCaMK), and it has recently been shown that, analogous to the cytokinin receptor, activation of CCaMK is sufficient to induce nodule formation in the absence of rhizobial bacteria (*8*, *9*). Hence, the spontaneous formation of nodules can be induced by gain-of-function mutations in both the *CCaMK* and *LHK1* genes. Tirichine *et al.* show that a plant carrying the gain-of-function mutations in both of these genes generates more nodules than plants with either mutation alone, indicating an additive effect. Hence, it is likely that two separate signaling pathways function in nodulation.

It is therefore very surprising that the *LHK1* gain-of-function mutation depends on the gene *NSP2* to activate spontaneous nodules. NSP2, a transcriptional regulator, is a component of the Nod factor signaling pathway (*10*, *11*) and is also necessary for CCaMK-induced spontaneous nodulation (*9*). It appears that even though these two pathways function independently, they converge at NSP2, suggesting that this regulator has dual functions in both Nod factor and cytokinin signal transduction during nodulation.

Perception of Nod factor by the plant is one of the first steps during the interaction between root cells and rhizobial bacteria. Because Nod factor accumulates in cell walls (*12*), it is highly unlikely that it can traverse the epidermis to induce responses in root cortical cells. It is possible that the localized production of cytokinins follows Nod factor perception, acting as a mechanism to coordinate epidermal and cortical responses during nodule formation. In this model, activation of Nod factor signaling at the epidermis leads to increased localized production of cytokinins. Cytokinins are then transported (most likely by active transport from cell to cell, but the mechanism remains unclear) to cortical cells where they are perceived by LHK1 at the cell surface. This initiates cell division, leading to formation of the nodule primordium. This model predicts that rhizobial bacteria induce cytokinin production in legume roots or redirect cytokinin transport, and also predicts that CCaMK-induced spontaneous nodulation should require LHK1.

Cytokinins are also used by plants that do not form nodules in diverse developmental pathways, including the regulation of root and shoot branching. A key question is whether plants that form nodules have evolved a unique response to cytokinins, or whether the appropriate activation of LHK1 and its orthologs in other plant species is sufficient to induce nodule-like structures. Answering this question will provide insights into the ease of transfer of nodule organogenesis, a first step in transferring this symbiotic interaction into agriculturally important species.

References

1. J. D. Murray *et al.*, *Science* **315**, 101 (2007); published online 16 November 2006 (10.1126/science.1132514).
2. L. Tirichine *et al.*, *Science* **315**, 104 (2007); published online 16 November 2006 (10.1126/science.1132397).
3. J. B. Cooper, S. R. Long, *Plant Cell* **6**, 215 (1994).
4. D. P. Lohar *et al.*, *Plant J.* **38**, 203 (2004).
5. S. Gonzalez-Rizzo, M. Crespi, F. Frugier, *Plant Cell* **18**, 2680 (2006).
6. U. Mathesius, C. Charon, B. G. Rolfe, A. Kondorosi, M. Crespi, *Mol. Plant Microbe Interact.* **13**, 617 (2000).
7. Y. Fang, A. M. Hirsch, *Plant Physiol.* **116**, 53 (1998).
8. L. Tirichine *et al.*, *Nature* **441**, 1153 (2006).
9. C. Gleason *et al.*, *Nature* **441**, 1149 (2006).
10. G. E. Oldroyd, S. R. Long, *Plant Physiol.* **131**, 1027 (2003).
11. P. Kalo *et al.*, *Science* **308**, 1786 (2005).
12. J. Goedhart, M. A. Hink, A. J. Visser, T. Bisseling, T. W. Gadella Jr., *Plant J.* **21**, 109 (2000).

10.1126/science.1137588

ECOLOGY

Rangeland Ecology in a Changing World

A synthesis of differnt models for assessing the balance between vegetation and livestock numbers in grazing lands may help policy-makers integrate human and ecosystem needs.

Lindsey Gillson and M. Timm Hoffman

Millions of people in rangelands depend directly on livestock for their livelihoods, but the management of these regions is mired in controversy (see the figure). Early colonial efforts imposed economic management models based on equilibrium carrying capacities. In the 1990s, however, "the new ecology" took hold and challenged equilibrium models by exploring the variability and flux inherent in most ecological and social systems (*1*). Such models (*2*) described a disequilibrium rangeland system, in which animal numbers were limited by drought events, thereby preventing an impact on vegetation (*3*). Since then, a variety of rangeland responses have been described, in which environmental variation and density-dependent effects on animal numbers are observed to a greater or lesser extent.

According to disequilibrium theory, herbivores have little impact on rangeland vegetation, because their numbers are regulated by environmental fluctuations (*2*). Droughts keep animal populations below levels where density-dependent competition for forage occurs (*3*). However, two problems have emerged from the disequilibrium literature. First, the process by which environmental variation regulates animal numbers is not clear. Second, if changes in vegetation are a normal part of variable, semiarid environments, then on what basis can variability be distinguished from potentially long-term degradation?

Recently, a new synthesis has addressed these challenges. It assumes that most rangeland systems respond to variability in rainfall. A fixed "carrying capacity," in the strict sense of a stable equilibrium between primary production and grazing, is unlikely because rainfall and hence primary production covary over time, particularly in semiarid ecosystems. If primary productivity varies with rainfall, then it follows that "carrying capacity" will also vary over time. Instead of stability, a correlation between animal numbers and rainfall might be expected (*4*), and indeed has been observed over long time scales (*5*).

Disequilibrium theory cannot explain this correlation because a central tenet is that animal numbers are below the level at which density-dependent competition for forage occurs (*2*, *3*). In low-rainfall years, however, density-dependent competition is possible even at quite low animal densities, because primary productivity is limited by lack of water, and animals will compete for the little forage that is available. Over time, no stable equilibrium between animal numbers and primary produc-

The authors are with the Institute for Plant Conservation, Botany Department, University of Cape Town, Rondebosch 7701, South Africa. E-mail: lindsey.gillson@uct.ac.za

tion will be reached because the position of central tendency changes according to rainfall, and the system can be described as nonequilibrium. At other times or scales, however, other environmental factors like drought, fire, nutrient limitation, forage quality, or rainfall frequency might dominate (*4*, *6*, *7*), and the coupling between rainfall and animal numbers will be weak.

Thus, a synthesis of disequilibrium and equilibrium theory describes a system that is not at equilibrium, but in which animal numbers fluctuate in response to a variable environment. At certain times, for example, during a series of low- to medium-rainfall years, or at key resource areas like dry-season grazing, these fluctuations may be mediated by density-dependent effects (*8*). At others, for example, during extreme environmental events like droughts, or outbreaks of disease, the system will be stochastically driven, and the correlation between animal numbers and rainfall will be weak. It is likely that disequilibrium, in its strict sense of a system dominated by environmental variation, can occur only in extremely dry environments and/or in exceptionally prolonged drought periods, when there is literally no primary productivity for which animals can compete (*9*). It therefore seems critical to distinguish disequilibrium systems (*3*) as a restricted subset of nonequilibrium systems, the latter being defined more broadly in terms of all systems that are not at equilibrium.

Given this broader definition of nonequilibrium, research into systems not at equilibrium can be reframed in terms of how factors like density dependence, environmental stochasticity, drought, disease, and nutrients interact; the thresholds at which different factors become dominant; and how these interactions are expressed in terms of plant-herbivore dynamics. Research has shown that consistently high stock levels can transform the vegetation composition of semiarid rangelands. For example, in Namaqualand, South Africa, a mixture of palatable woody plants and perennial herbaceous species was replaced through heavy grazing by a system dominated by unpalatable woody species and annuals (*10*). In the Rio Grande Plain, United States, and New South Wales, Australia, grass-dominated rangelands were replaced by a more wooded landscape when grazing levels were high and fire was suppressed (*11*, *12*), whereas in Kimberley, South Africa, rainfall frequency and nutrient availability were found to be more important factors than animal abundance in woody-plant encroachment (*13*). Rather than a continuum of variation in vegetation, then, rapid transitions in vegetation types have been observed when the interactions among livestock density, rainfall, grazing history, fire, and nutrients reach critical thresholds (*14*, *15*).

Sensitive to environmental conditions. Millions of people in drylands depend directly on livestock for their livelihoods.

Systems theory can help us understand rangeland complexity. The concept of carrying capacity—long associated with assumptions of stability and equilibrium—can be usefully replaced by the idea from chaos theory of a "moving attractor" (*11*, *16*). In rangelands, this would be a point that varies depending on rainfall and toward which animal numbers tend. Similarly, the state-and-transition understanding of rangelands (*14*) can be reinterpreted in terms of the moving attractor concept, where alternative persistent states may be considered as regions of higher-probability space, known in systems theory as a domain of attraction.

Conceptual frameworks like stable states and moving attractors may also provide a useful underpinning for addressing the second major research challenge, that of how to define degradation in a variable system. Annual-dominated systems might maintain a long-term average in animal numbers and therefore not be degraded in terms of animal productivity, but might still be considered degraded in terms of the loss of plant diversity and cover that can accompany heavy grazing, or because of effects on soil erosion or nutrient cycling. Moreover, a grazing system dependent on annuals is likely to become disequilibrial, that is, driven by environmental stochasticity, because of the increased likelihood of a complete failure of primary productivity in drought years.

Analysis of vegetation dynamics (*4*), clarity in the definition of degradation (*17*), and a refinement of disequilibrium theory to incorporate density dependence at different scales, as well as environmental stochasticity and variability in forage quality (*7*), have opened new opportunities for dialogue between the natural and social sciences, particularly with regard to the interplay between ecological, social, and economic systems (*4*, *18*). However, this still leaves policy-makers with the challenge of how to integrate the views of multiple stakeholders, with different value systems, in order to reconcile biodiversity conservation and livelihood concerns across contested rangeland landscapes.

References and Notes

1. I. Scoones, *Annu. Rev. Anthropol.* **28**, 479 (1999).
2. J. E. Ellis, D. M. Swift, *J. Range Manag.* **41**, 450 (1988).
3. R. H. Behnke, I. Scoones, C. Kerven, Eds. *Range Ecology at Disequilibrium: New Models of Natural Variability and Pastoral Adaptation in Africa Savannas* (Overseas Development Institute, London, UK, 1993).
4. B. M. Campbell *et al.*, *Ecol. Econ.* **60**, 75 (2006).
5. T. A. Benjaminsen *et al.*, *Ann. Assoc. Am. Geogr.* **96**, 224 (2006).
6. S. Vetter, *J. Arid Environ.* **62**, 321 (2005).
7. D. D. Briske, S. D. Fuhlendorf, F. E. Smeins, *J. Appl. Ecol.* **40**, 601 (2003).
8. A. W. Illius, T. G. O'Connor, *Ecol. Appl.* **9**, 798 (1999).
9. S. Sullivan, R. F. Rohde, *J. Biogeogr.* **29**, 1 (2002).
10. F. M. Chambers, D. Mauquoy, P. A. Todd, *J. Appl. Ecol.* **36**, 719 (1999).
11. M. A. Janssen, J. M. Anderies, B. H. Walker, J. *Environ. Econ. Manag.* **47**, 140 (2004).
12. B.-L. Li, *Acta Biotheoretica* **50**, 141 (2002).
13. W. J. Bond, F. I. Woodward, G. F. Midgley, *New Phytol.* **165**, 525 (2005).
14. M. Westoby, B. Walker, I. Noy-Meir, *J. Range Manag.* **42**, 266 (1989).
15. A. M. Cingolani, I. Noy-Meir, S. Díaz, *Ecol. Appl.* **15**, 757 (2005).
16. A. Hastings *et al.*, *Annu. Rev. Ecol. Syst.* **24**, 1 (1993).
17. J. F. Reynolds, D. M. Stafford Smith, in *Global Desertification: Do Humans Cause Deserts?*, J. F. Reynolds, D. M. Stafford Smith, Eds. (Dahlem Univ. Press, Berlin, 2002), pp. 1–22.
18. J. E. Gross *et al.*, *Environ. Model. Software* **21**, 1264 (2006).

19. We thank N. Allsopp, D. Brockington, P. Carrick, R. Rohde, and S. Vetter for helpful comments on the text.

10.1126/science.1136577

CREDIT: KRISTINA DUFFIN

INTRODUCTION

Catching Cosmic Clues

THE AUTHOR HENRY JAMES WROTE THAT "EXPERIENCE IS ... A KIND OF HUGE spider-web of the finest silken threads suspended in the chamber of consciousness, and catching every airborne particle in its tissue." Particle astrophysicists are trying to weave their own webs by building vast detectors on Earth and in space that will ensnare cosmic particles and so teach us about the building blocks of the universe.

Thanks to enormous progress in cosmology in recent years, astrophysicists are both pleased and perplexed. On the one hand, they have succeeded in nailing down the universe's mass, geometry, and expansion rate. But on the other, they have discovered that 95% of the stuff of the universe is in two unknown forms that they have named "dark matter" and "dark energy." Only 5% is normal matter: electrons, protons, and neutrons. Pinning down the nature of this missing mass and energy is difficult, because dark matter does not absorb light or interact with normal atoms; the dark energy driving accelerated cosmic expansion is even more intangible. Particle physicists may, however, have the tools to test some ideas. In this special issue devoted to particle astrophysics, a rapidly developing interdisciplinary area, six Perspectives cover not only candidates for dark matter but also the physics of the Big Bang fireball, neutrinos, cosmic rays, and sources of extreme-energy gamma rays such as black holes.

Neutrino physics has leapt ahead in recent years, with measurements of neutrino mass and oscillations between different types, or flavors. The next frontier is neutrino astronomy, capturing neutrinos from sources more distant than the Sun, and vast arrays of detectors are being built under the ice in Antarctica and under the Mediterranean Sea to do this. Neutrinos hardly interact with normal matter at all, but occasionally they do and produce ghostly flashes of light that detectors can catch. If the universe's hidden mass takes the form of other particles, then axions and WIMPs (weakly interacting massive particles) are the prime suspects. Experiments, many hidden below ground to isolate the detectors from other stray particles, have been designed and are being implemented to spot these exotic particles via their recoil off other nuclei. Currently, these detectors are modest in size, but detectors now on the drawing board could weigh as much as a ton.

High-energy particles can also be used for astronomy. Cosmic-ray observatories are nearing the sensitivities required to detect individual sources in the sky, thus testing acceleration physics. Cosmic rays are created by extreme astrophysical sources such as supernova shock waves, gamma-ray bursts, and near black holes. Very-high-energy gamma-ray emission from these sources is already detectable with new telescope arrays and has constrained the physics of particle jets emanating from compact stars and black holes.

Particle astrophysics is an exciting area brimming with promise. As scientists come together to combine their know-how, maybe in the next decade we will find the missing matter, and crown the already remarkable achievements of cosmology.

– JOANNE BAKER

CREDIT: NSF

Particle Astrophysics

CONTENTS

Science

NEWS

Stalking Discovery From the Infinitesimal to the Infinite

Particle physicists are moving into astrophysics, astronomy, and cosmology; their skills and big-hammer approach could help solve some of the universe's deepest mysteries

STRETCHING FOR HUNDREDS OF KILOMETERS and covered with scrub and prairie grass, the Pampa Amarilla in western Argentina would be an ideal place to graze cattle or film a Western or, on a clear night, gaze at the stars and contemplate one's place in the cosmos. But James Cronin, a particle physicist at the University of Chicago in Illinois, has chosen this unlikely venue to try to solve an enduring mystery of astrophysics.

Cronin and 300 colleagues have come to the foot of the Andes mountains to snare particles from deep space that zing along with energies millions of times higher than particle accelerators have achieved on Earth. If all works as hoped, in a few years researchers will spot the sources from which such cosmic rays emanate. "That's never been done, and that would be a huge breakthrough," says Cronin, who shared the Nobel Prize in physics in 1980 for the discovery of a slight asymmetry between matter and antimatter known as CP violation.

The experiment is no small undertaking. Researchers are carpeting the plain with 1600 detectors spaced 1.5 kilometers apart to sense the avalanche of particles created when a ray crashes into the atmosphere. When it is completed, the Pierre Auger Observatory will cover 3000 square kilometers—five times the area of Chicago. True to his particle physicist's training, Cronin embraces a simple credo: "Just think big."

Cronin is only one of many particle physicists who are turning away from Earth-bound accelerators and toward the heavens. In recent years, researchers have begun explorations at the boundaries between particle physics, astrophysics, and astronomy. They are lurking in caves trying to detect particles of the dark matter that holds the galaxies together; sinking detectors into the ice at the South Pole and the waters of the Mediterranean Sea to sense particles called neutrinos from outer space; building gamma ray telescopes to open new eyes on the cosmos; and tracking stellar explosions known as supernovae to decipher the space-stretching dark energy that is accelerating the expansion of the universe. All these endeavors fall under the nebulous rubric of particle astrophysics, or astroparticle physics.

"It's likely that in the next 10 years, one of these efforts will lead to a major discovery," says Gerard van der Steenhoven, a particle physicist at the National Institute for Nuclear and High Energy Physics in Amsterdam, the Netherlands, who works on a neutrino experiment in the Mediterranean. "That makes it very exciting."

The growth of particle astrophysics is not only rejuvenating particle physics but also changing astrophysics and astronomy. Accustomed to working on immense experiments in huge collaborations, particle physicists bring their skills and strategies to fields in which the experiments are already growing rapidly in size and complexity. "You're bringing in a new culture and a new way of operating at a time when the field [of astronomy and astrophysics] needs it," says Bruce Winstein, a particle physicist at the University of Chicago who now studies the

On the range. Physicists with the Pierre Auger Observatory are covering 3000 square kilometers of Argentine prairie with particle detectors.

CREDIT: PIERRE AUGER OBSERVATORY

afterglow of the big bang, the cosmic microwave background radiation.

But whether particle astrophysics continues to flourish may depend on whether experiments currently in the works deliver any of the hoped-for discoveries. In fact, some say, the future of the field could depend in part on what researchers find at the next great particle collider, currently under construction in Europe.

Cosmic connections

In turning toward astrophysics, particle physics is, in a sense, returning to its roots. Physicists spotted the first bit of antimatter—the antielectron, or "positron"—while studying cosmic rays in 1932. In the same way, they discovered the first particle beyond those that make up the everyday matter around us, the muon, a few years later. But particle astrophysics stretches beyond the study of particles from space. It represents a broad movement of particle physicists into fields such as cosmology and astronomy, where they are pursuing the grandest mysteries in the universe, sometimes without a particle in sight.

Most physicists trace the growth of the field to conceptual links between cosmology and particle physics forged in the 1970s and 1980s. For example, theorists realized that the abundance of helium in the universe puts a limit on the number of possible types of neutrinos, wispy particles produced in certain kinds of radioactive decay that interact feebly with everyday matter. (Physicists now know that there are three types of neutrinos.) Others noted that, when mixed into the big bang theory, CP violation might explain why the universe contains so much matter and so little antimatter.

Still others realized that a particle theory might help explain the nature of dark matter, the unidentified stuff whose gravity holds the galaxies together. The standard model of particle physics says that matter is made of particles called quarks and leptons that exchange force particles called bosons. A theory called supersymmetry extends this scheme by positing that every known fundamental particle has a more massive doppelgänger that has yet to be discovered. Some of those particles might just fit the bill for dark matter.

Such connections have blurred the distinction between particle physics and cosmology, says Jonathan Ellis, a theorist at the European particle physics laboratory CERN near Geneva, Switzerland. "I often find it difficult to tell when I'm writing a paper on particle physics and when I'm writing a paper on cosmology, because in my mind the two are inextricably intertwined," he says.

CREDIT: JOHN JACOBSEN/NSF

Brrr! Researchers constructing Ice Cube lower a photodetector into the South Pole ice.

More recently, experimenters have joined the movement to particle astrophysics, inspired by key discoveries made in recent years. Closest to home, the biggest advance in particle physics in the past 2 decades came from researchers studying neutrinos from space with the Super-Kamiokande particle detector in a mine in Japan. In 1998, physicists found that one type of neutrino could transform into another, a phenomenon known as mixing that can occur only if neutrinos have mass. The standard model assumes that neutrinos are massless, so the observation gives researchers their first peek at physics beyond the standard model.

Further afield, scientists studied distant stellar explosions known as type Ia supernovae to trace the history of the expansion of the universe. In 1998, two groups independently reported that the most distant supernovae were even farther away than expected, indicating that the expansion of the universe is accelerating. That stunning observation suggested that some mysterious "dark energy" is stretching the fabric of space.

That revolutionary notion was bolstered in 2003 when researchers working with NASA's Wilkinson Microwave Anisotropy Probe satellite mapped the cosmic microwave background in exquisite detail. Analyzing the tiny temperature differences in the radiation across the sky, they found that the universe consists of roughly 71% dark energy, 24% dark matter, and just 5% ordinary matter.

The very notions of dark energy and dark matter fire the imaginations of researchers who have devoted themselves to asking, "What's it made of?" says Natalie Roe, a particle physicist at Lawrence Berkeley National Laboratory (LBNL) in California. "Having realized that quarks and leptons are only 5% of the universe, I think it's only natural to ask what the other 95% is," she says. "So dark energy and dark matter are natural targets for particle physicists."

Making a move

When explaining their switch into particle astrophysics, researchers cite motives as varied as the particles in the standard model. Most say they were drawn by the intellectual excitement of a young field. "Particle physics was most exciting before the standard model was put in final form and verified," says Steven Weinberg, a theorist at the University of Texas, Austin, who shared the Nobel Prize in physics in 1979 for his work on the standard model and now pursues cosmology. "In cosmology, the questions are more wide open."

LBNL's Roe, who spent a decade studying the properties of quarks to high precision, says she finds it refreshing to work in a field in which researchers generally don't know what to expect from an experiment. "I wanted to look into something that we really didn't understand, where we don't have a standard model," says Roe, who is working on a satellite, the Supernova/Acceleration Probe, that would examine dark energy by measuring thousands of supernovae.

Many researchers say they switched to particle astrophysics in search of a more congenial work environment. Daniel Akerib, a particle physicist at Case Western Reserve University in Cleveland, Ohio, says he moved away from collider experiments, which typically involve hundreds of collaborators, so he could take a more hands-on approach to his work. "I just felt like I was going to spend all my time in meetings and not have any fun," he says. Akerib now works with the Cryogenic Dark Matter Search (CDMS), a small group that runs an extremely sensitive detector in a mine in Minnesota and hopes to spot passing dark-matter particles.

Some physicists have set out in new directions because opportunities in particle physics have dwindled. David Cinabro of Wayne State University in Detroit, Michigan, had been working on an experiment called BTeV that would have run at the Tevatron collider at Fermi National Accelerator Laboratory (Fermilab) in Batavia, Illinois. But in 2005, the U.S. Department of Energy suddenly axed the project. "I

was faced with the prospect of starting over no matter what I did," Cinabro says.

Cinabro could have joined one of the experiments at the next great accelerator, the Large Hadron Collider (LHC) at CERN, which is scheduled to turn on late this year. Instead, he joined the Sloan Digital Sky Survey, a novel astronomy effort that uses a 2.5-meter optical telescope on Apache Point, New Mexico, to map everything in one quadrant of the sky. Making the shift wasn't easy, says Cinabro, who is studying supernovae and dark energy. "It's like going back to graduate school, because I'm as ignorant as a first-year graduate student," he says. Still, he says he's happy with his decision.

Homey. Experimenters can take a more hands-on approach with the relatively small CDMS dark-matter detector.

A few researchers say they have pursued particle astrophysics for the sheer adventure of it. "To me it was an opportunity to see Antarctica through the back door and not have to pay for it," quips David Besson of the University of Kansas, Lawrence, who is working on a prototype neutrino detector at the South Pole. In a phone interview from McMurdo Station, Besson says there is something romantic about searching for radio signals produced by cosmic neutrinos crashing into the ice. "It takes you back to that sense of wonder when you were 5 years old and you'd look up and see the stars," he says. "Not that you could do that where I grew up in New Jersey."

Rearranging the furniture

As interest in particle astrophysics has grown, so has funding for such research. For example, in 2000, the U.S. National Science Foundation (NSF) instituted a program in particle and nuclear astrophysics, which now has a $16 million annual budget. And since 1994, the United Kingdom has funded particle physics, particle astrophysics, and astronomy jointly out of its Particle Physics and Astronomy Research Council. But even as the growth of particle astrophysics is expanding the boundaries of particle physics, it is also changing the practice of astronomy and astrophysics.

Most obviously, particle physicists bring with them technologies that are opening new avenues of inquiry. For example, NASA's Gamma-Ray Large Area Space Telescope (GLAST), which is scheduled for launch this October, will provide astronomers with an unparalleled view of the universe as seen in very-high-energy photons. But the "camera" that will detect the gamma rays is a particle detector built at the Stanford Linear Accelerator Center (SLAC) in Menlo Park, California. "In the end, what you need is the best equipment you can get to solve the problem. And if it comes from some other field, why not?" says SLAC's Eduardo do Couto e Silva.

Particle physicists have also introduced a different style of collaboration to astronomy and astrophysics, as exemplified by the Sloan survey. When sharing a telescope, astronomers traditionally allot observers time to use the instrument in turn. In contrast, Sloan researchers pull together to crank out a steady stream of data in a general format, so that collaborators can analyze the data any way they please, just as in a collider experiment. In essence, the Sloan telescope produces astronomical data just as a factory might produce brake pads.

Perhaps most important, particle physicists have appetites for huge projects that push the limits of technology, organization, and funding. "These are not people who are afraid to ask for big things, and they're used to people saying yes," says Michael Turner, a cosmologist at the University of Chicago who served as assistant director of NSF's mathematics and physical sciences directorate from October 2003 until April 2006. As particle physicists enter astrophysics and astronomy, their habit of "thinking big" is accelerating the natural growth of the size of projects, Turner says.

But even as particle astrophysics blossoms, some researchers worry about its future. Steven Ritz, a particle physicist at NASA's Goddard Space Flight Center in Greenbelt, Maryland, and project scientist for GLAST, fears that the rise of particle astrophysics could undermine accelerator-based research. "Sometimes the movement is interpreted to mean that there's no need to build accelerators anymore, that you can do it all from space," he says, "and that's just not right." Even so, the number of colliders is falling, especially in the United States. SLAC will shutter its PEP-II collider in 2008, and a year later Fermilab will unplug the Tevatron, leaving the United States with no colliders for particle physics.

Others say the growth of particle astrophysics will likely slow as the size and expense of projects balloons. "It will soon hit a wall that particle physics hit some time ago, and that is the $1 billion experiment," says Francis Halzen, a particle theorist–turned-experimenter at the University of Wisconsin, Madison. Halzen's own experiment, Ice Cube, exemplifies the growth of projects in particle astrophysics. A mammoth array of photodetectors being embedded between 1.5 and 2.5 kilometers deep in the South Pole ice, Ice Cube will detect light produced when ultrahigh-energy neutrinos crash into the ice. Scheduled for completion in 2011, the experiment will cost $271 million and involve 400 researchers.

Most of all, the future of particle astrophysics depends on what experiments currently in the works might find. Roger Blandford, a theoretical astrophysicist at Stanford University in Palo Alto, California, says the first big test will come in the search for dark matter. "Our working hypothesis is that dark matter comprises supersymmetric particles," he says. "We could be terribly wrong." Given that hypothesis, the prospects for the searches would brighten if the LHC discovers supersymmetric particles—and dim if it doesn't.

For the moment, researchers working in particle astrophysics are happy just to participate in such a young and dynamic field. Promises of momentous discoveries abound. Expectations are sky high.

–ADRIAN CHO

CREDIT: REIDAR HAHN/FERMILAB VISUAL MEDIA SERVICES

PERSPECTIVE

Quarks and the Cosmos

Michael S. Turner

Cosmology is in the midst of a period of revolutionary discovery, propelled by bold ideas from particle physics and by technological advances from gigapixel charge-coupled device cameras to peta-scale computing. The basic features of the universe have now been determined: It is 13.7 billion years old, spatially flat, and expanding at an accelerating rate; it is composed of atoms (4%), exotic dark matter (20%), and dark energy (76%); and there is evidence that galaxies and other structures were seeded by quantum fluctuations. Although we know much about the universe, we understand far less. Poised to dramatically advance our understanding of both the universe and the laws that govern it, cosmology is on the verge of a golden age.

The universe is often just beyond our grasp, and progress in cosmology usually comes only with advances in technology or with powerful new ideas. In the 1920s, Hubble used the new 100-inch Hooker telescope on Mount Wilson to discover the expansion of the universe, and Einstein's young theory of general relativity provided the mathematics needed to understand our Big Bang beginning. Only with the advent of the 200-inch Hale telescope on Mount Palomar in the 1960s did astronomers push to the edge of the observable universe, and radio technology made possible the discovery of the microwave echo of the Big Bang in 1964 by Arno Penzias and Robert Wilson, revealing that, at its creation, the universe was hot as well as dense.

Cosmology slipped into the doldrums in the 1970s and was aptly described by Sandage as the search for two numbers: the expansion rate H_0 and the deceleration parameter q_0 (*1*). Although the basics of the hot Big Bang model were in place, including the picture of how structure in the universe formed by gravity amplifying small variations in the matter density into galaxies, clusters of galaxies, and superclusters, there was no evidence for the tiny seed inhomogeneities that were required to form structure or hints as to their origin. Moreover, our understanding of cosmology hit a brick wall at 10^{-5} s, which blocked connecting with its earliest moments and a deeper understanding of the Big Bang. Because of missing physics, the early universe was thought to be a confusing sea of overlapping protons, neutrons, and other elementary particles.

Today, cosmology is in the midst of a revolutionary period of discovery. In the past 8 years, the field has twice captured *Science*'s Breakthrough of the Year: for the discovery of the acceleration of the expansion of the universe in 1998 (*2*) and for the concordance cosmological model in 2003 (*3*). The revolution traces its beginnings to the 1980s with the arrival of powerful new ideas and advances in technology: from the discovery of quarks and the introduction of charge-coupled devices to space-based telescopes and string theory.

New ideas from particle physics changed the language as well as the conversation in cosmology. Physically based quantities climbed to the top of the list of wanted parameters: the temperature, spectrum, and anisotropy of the cosmic microwave background (CMB); the shape of the universe; the composition of the universe; the large-scale distribution of matter today; and the spectrum of seed inhomogeneities.

By using bigger telescopes, better detectors, and faster computers, astronomers and physicists have determined all of these parameters and more to percent-level precision, turning an oxymoron—precision cosmology—into reality. And in most cases, it is not just one measurement but an interlocking web of complementary determinations that pin down the parameters, strengthening the framework and changing the tenor of cosmology. Cosmology is no longer the field described by the Russian physicist Lev Landau who said, "Cosmologists are often in error, but never in doubt."

This then is our universe: On the whole, it is spatially flat and 13.7 billion years old, both of which are known to 1% precision; it is expanding at a rate of 70 ± 2 km/s per megaparsec, and the expansion is speeding up; and it is composed of $24 \pm 4\%$ matter and $76 \pm 4\%$ dark energy, with $4.2 \pm 0.5\%$ of the matter in the form of atoms, between 0.1 and 1% in the form of neutrinos, and with the bulk of the matter dark and as yet unidentified (*4*). Stars, the only constituent of Sandage's universe, account for less than 1% of the total composition. The microwave background temperature has been measured to four significant figures, 2.725 ± 0.001 K (*5*), and its tiny variations (about 0.001%) across the sky have been mapped with a resolution of better than 0.1° (*6*).

The discovery of quarks (the constituents of neutrons and protons) and the realization that they interact weakly when close together knocked down the brick wall and opened the door to understanding the nature of the very early universe: It was a hot soup of quarks and other elementary particles and almost as easy to describe as the chemists' perfect gas. Thinking of the early universe as quark soup changed the big questions: Where is the antimatter? What is the origin of the seed inhomogeneity? Why is the universe so smooth, nearly flat, and very old? Where did the heat of the Big Bang—today existing in the billion CMB photons per atom—come from? What powered the Big Bang?

Speculations about the earliest moments of creation and possible answers to all the big questions, based on bold ideas about how the fundamental particles and forces of nature are unified, burst forth. Many ideas have been influential (for example, how neutrino mass and "charge-parity" violation can explain the absence of antimatter and the few atoms per billion photons in the universe today), but two ideas have been central to the current revolution: dark matter as a new form of matter, and inflation as a dynamical explanation for the most salient features of the universe.

As cosmological observations were establishing that there was insufficient atomic matter to account for the vast amounts of dark matter needed to hold together cosmic structures from galaxies to superclusters, particle physics came forward with three well-motivated candidates for dark matter. The fate of the first candidate, the neutrino, turned on neutrino mass; we now know that neutrinos have mass and are part of the dark matter, but only a tiny part of it, accounting for less than 1%. Hopes are now pinned on two still-to-be-discovered particles: (i) the neutralino, which is expected to have a mass of about 100 times that of the proton and to be the lightest of a new class of particles predicted by string theory, and (ii) the axion, a particle that is expected to be a trillion times less massive than the electron (*7*).

The central tenet of inflation is a very early burst of accelerated expansion driven by yet-to-be-understood physics involving a new scalar field called the inflaton. This rapid expansion led to a smooth, flat, and hot universe (the heat produced by the conversion of the inflaton's energy into particles), and quantum fluctuations, blown up from subatomic scales to astronomical scales, created the density inhomogeneities that seeded all cosmic structure (*8*).

Taken together, inflation and dark matter led to the cold dark matter (CDM) theory of structure formation (cold refers to the fact that the dark matter particles such as the axion and neutralino move slowly, which leads to predictions that do not depend on the mass of the dark matter particle) (*9*). CDM describes in detail how cosmic structure formed and how the bright side of the universe came to be, and it has stimulated the observations (made possible by new technology) that have now filled in the story line of the formation and evolution of galaxies and large-scale structure, from shortly after the birth of the first stars (less than a billion years after the Big Bang) until today.

NASA's Cosmic Background Explorer (COBE) satellite, which first detected the tiny variations in the CMB intensity across the sky (anisotropy) in 1992 (*10*), confirmed the existence of the underlying matter inhomongeneity that seeded structure and began a new era in

Kavli Institute of Cosmological Physics, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637–1433, USA. E-mail: mturner@uchicago.edu

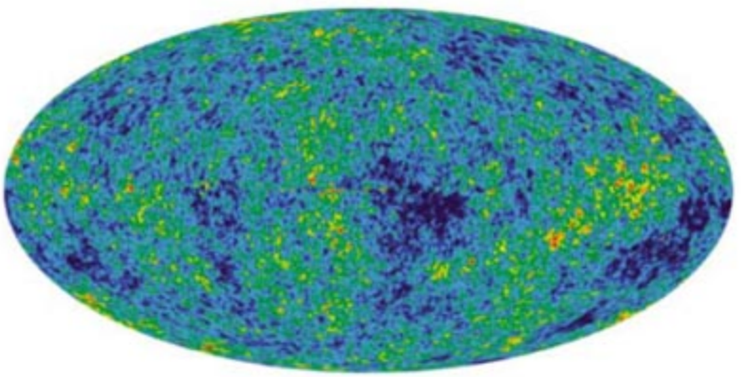
cosmology. Encoded in the CMB anisotropy is information about the past, present, and future of the universe (Fig. 1). (In 2006, John Mather and George Smoot received the Nobel Prize in Physics for their work on COBE.)

COBE was followed up by NASA's Wilkinson Microwave Anisotropy Probe (WMAP) and a host of ground- and balloon-based CMB experiments. The results from these higher-resolution and higher-precision measurements, together with maps of the large-scale structure of the universe made by the Sloan Digital Sky Survey and Two-Degree Field project, have precisely determined cosmological parameters, established the basic correctness of the CDM paradigm, and provided strong support for the ideas of particle dark matter and inflation (*4*).

Inflation and particle dark matter were new ideas by design; dark energy, on the other hand, came as a surprise. The quest for Sandage's second number, q_0, took an unexpected turn in 1998. Armed with new technology and a better standard candle to determine cosmic distances (type Ia supernovae, the nuclear explosions associated with white dwarf stars pushed over the Chandraskehar mass limit by accretion from a companion), two teams presented evidence that the expansion of the universe is speeding up, not slowing down (*11*, *12*): that is, q_0, so carefully defined to be positive, is actually negative!

Although the mystery of cosmic acceleration surely ranks as one of the most profound puzzles in all of science today, it was also the missing piece that pulled the current picture together. Toward the end of 1990s, the inflation/CDM paradigm was working well except for one "small" detail: There was growing evidence for both a flat, critical-density universe and for a matter density that was only 30% of the critical density. Where was the other 70% of the critical density? Cosmic acceleration solved the problem: The observed cosmic speed up indicated the existence of a very smooth and diffuse form of very elastic energy (now referred to as dark energy), which accounts for the missing 70%. When the discovery of cosmic acceleration came, the current concordance model, as absurd as it seems, was quickly embraced. [In fact, two theoretical papers anticipated this solution a few years before the discovery (*13*, *14*).]

Within general relativity, dark energy can account for cosmic acceleration because Einstein's theory predicts that a substance whose pressure is more negative than one-third of its energy density has repulsive gravity. Ideas about what dark energy is range from the energy of the quantum vacuum to the existence of another new scalar field (called quintessence and possibly related to the inflaton) to the influence of unseen additional spatial dimensions predicted by string theory (*15*). The most interesting possibility of all is the absence of dark energy, with cosmic acceleration being explained by a new aspect of gravity, one not accounted for by general relativity (*16*, *17*).



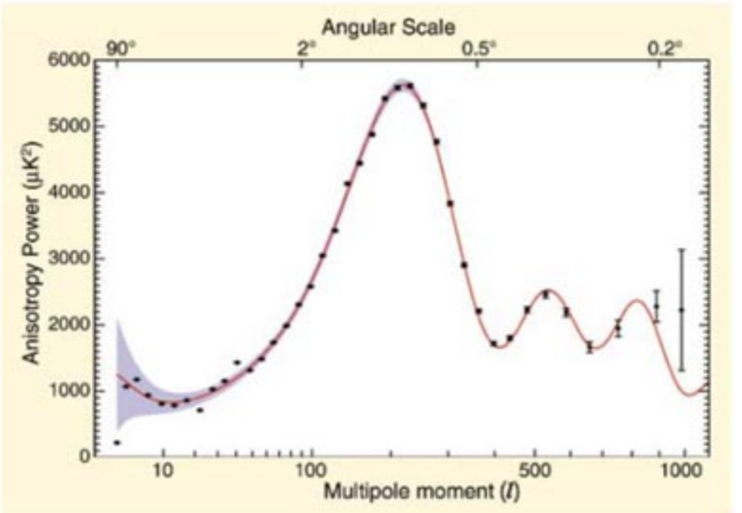


Fig. 1. The CMB seen by WMAP. (**Top**) The hot (red) and cold (blue) spots on microwave sky, as measured by WMAP, map out the tiny inhomogeneities in the distribution of matter when the universe was about 400,000 years old; the full range of the variations is ±200 μK, corresponding to variations of about 0.01% in the matter density. (**Bottom**) The spherical-harmonic multipole content of the anisotropy reveals the underlying mathematical structure that has been used to determine cosmological parameters and to provide for cosmic inflation. The points and their error bars indicate the WMAP measurements and their estimated errors, and the solid curve is the prediction of the concordance cosmological model. The shaded area indicates the "cosmic variance" interval for each multipole, which fundamentally limits the precision with which the underlying theory can be tested. (Cosmic variance arises because the underlying multipole distribution is being estimated by measuring the $2l + 1$ multipole moments that can be determined for a given l.) [Image: NASA/WMAP Science Team]

The roots of the dark-energy puzzle extend back to the birth of quantum mechanics and Einstein's famous fudge factor. According to quantum mechanics, the vacuum should be filled with a sea of "virtual" particles whose existence is allowed by the uncertainty principle. The effects of these virtual particles are very real (they shift atomic lines and elementary particle masses) and have been measured. With a bulk pressure equal to the negative of its energy density and mathematically equivalent to Einstein's cosmological constant, quantum vacuum energy would seem to be the obvious explanation for cosmic acceleration. However, there is one small problem: When theorists try to calculate how much quantum nothingness weighs, they get a number that is absurdly large (one that is actually infinite). The so-called cosmological constant problem (*18*), which has been around for more than 30 years, can no longer be ignored because it is now tied to understanding why the expansion of the universe is speeding up.

Today, we know much about the universe—its shape, age, composition, evidence for an inflationary beginning, and timeline from quark soup through the formation of large-scale structure (Fig. 2)—but we understand little about the universe with its odd mixture of atoms, dark matter, and dark energy. We do not know what the bulk of the dark matter is, why the expansion is accelerating today, or if the universe actually underwent an early burst of inflationary expansion and, if so, what caused it. To put a new twist on Landau's words, cosmologists today are rarely in error but are often in doubt.

With the new accelerators, telescopes, and experiments on the horizon, there is certainly much more to come in cosmology over the next 15 years. The neutralino could be produced at Fermilab's Tevatron or the Large Hadron Collider at CERN, or the neutralinos or axions that hold together the Milky Way could be detected by an ultrasensitive detector (*7*). A new generation of CMB experiments zeroing in on the polarization could reveal the third and most definitive signature of inflation (i.e., the gravitational waves that arise because of quantum fluctuations in the metric of spacetime) and pin down when inflation occurred. A dark-energy space telescope could shed light on why the universe is accelerating.

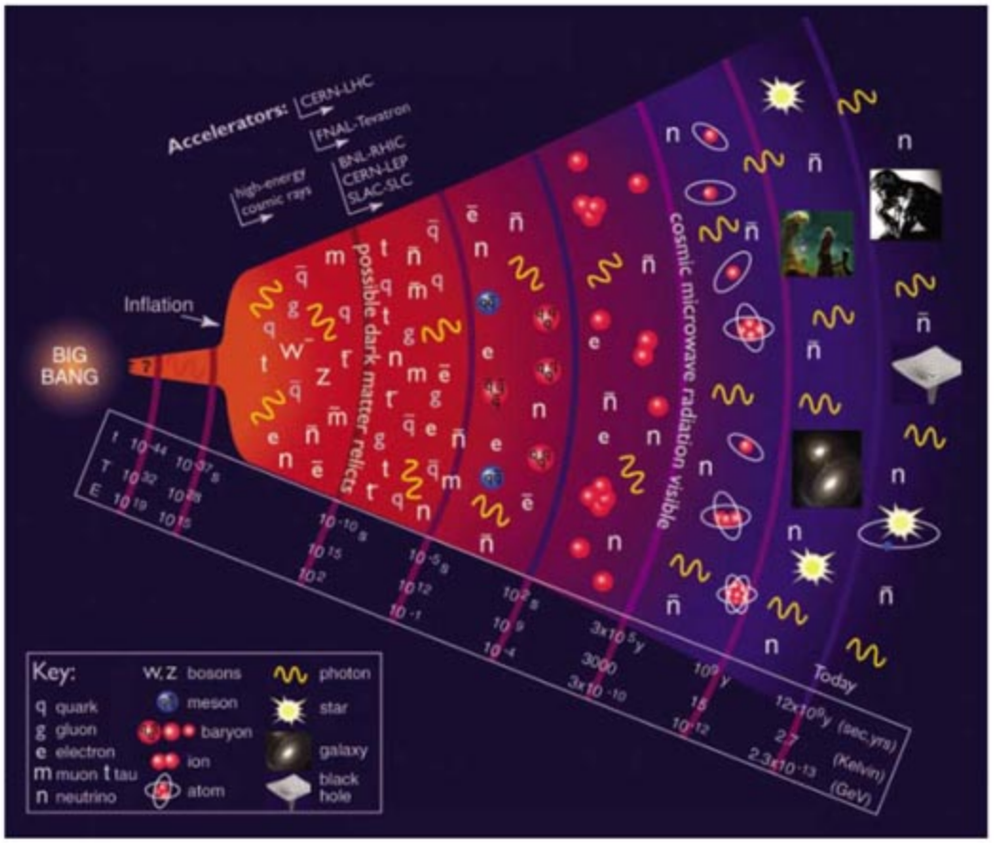


Fig. 2. The Big Bang timeline, from inflation to quark soup to the birth of the light nuclei to the formation of atoms and ultimately of galaxies and other gravitationally bound structures. [Image: Particle Data Group/Lawrence Berkeley National Laboratory/U.S. Department of Energy/NSF]

Because the scientific agendas of cosmology and particle physics have converged, as we deepen our understanding of the universe, we will advance our understanding of the fundamental laws that govern it as well. Nowhere in particle physics are the stakes higher than for string theory. If string theory is to live up to its billing as "the theory of everything" rather than, as some say, a theory of nothing, it needs a home run. Because most of its current predictions exceed the reach of terrestrial laboratories, many string theorists are pinning their hopes on a cosmological home run, such as a fundamental understanding of inflation (or a more attractive alternative), a solution to the puzzle of cosmic acceleration, or insight into the nature of the Big Bang itself.

Beyond the next 15 years, the future of cosmology is less clear. The universe could, as it has before, slip beyond our reach. Just as particle physicists were simultaneously blessed and cursed with the success of their standard model, we could find ourselves with even more precision but no more understanding. Nevertheless, I am bullish and predict that this longest boom in cosmology will ultimately earn the status of a golden age, for dramatically advancing the understanding of both quarks and the cosmos.

References and Notes

1. A. R. Sandage, *Phys. Today* **23**, 34 (1970).
2. J. Glanz, *Science* **282**, 2156 (1998).
3. C. Seife, *Science* **302**, 2038 (2003).
4. M. Tegmark *et al.*, preprint available at http://arxiv.org/abs/astro-ph?papernum=0608632.
5. J. Mather *et al.*, *Astrophys. J.* **512**, 511 (1999).
6. D. Scott, G. Smoot, preprint available at http://arxiv.org/abs/astro-ph?papernum=0601307.
7. B. Sadoulet, *Science* **315**, 61 (2007).
8. See E. W. Kolb, M. S. Turner, *The Early Universe* (Westview, New York, 1990), chap. 8.
9. G. Blumenthal *et al.*, *Nature* **311**, 517 (1984).
10. G. Smoot *et al.*, *Astrophys. J.* **396**, L1 (1992).
11. S. Perlmutter *et al.*, *Astrophys. J.* **517**, 565 (1999).
12. A. Riess *et al.*, *Astron. J.* **116**, 1009 (1998).
13. L. Krauss, M. S. Turner, *Gen. Relativ. Gravitat.* **27**, 1137 (1995).
14. J. P. Ostriker, P. J. Steinhardt, *Nature* **377**, 600 (1995).
15. E. J. Copeland, M. Sami, S. Tsujikawa, preprint available at http://arxiv.org/abs/hep-th?papernum=0603057.
16. S. Carroll *et al.*, *Phys. Rev. D* **70**, 043528 (2004).
17. C. Deffayet, G. Dvali, G. Gabadadze, *Phys. Rev. D* **65**, 044023 (2002).
18. S. Weinberg, *Rev. Mod. Phys.* **61**, 1 (1989).

10.1126/science.1136276

PERSPECTIVE

Particle Dark Matter in the Universe: At the Brink of Discovery?

Bernard Sadoulet

The nature of dark matter is one of the central problems of cosmology, particle physics, and gravity. It may be made of still unknown particles produced in the early universe. Much progress has been made in attempts to detect these particles and in the development of the required experimental techniques. Results from direct searches, the Large Hadron Collider, and the Gamma-ray Large Area Space Telescope offer promising opportunities within the next decade to find the missing dark matter.

The past decade of precision cosmological observations has led us to a surprising model of the universe (*1*, *2*): Ordinary matter (baryons and electrons) represents only 5% of its energy density; the rest does not interact with photons and constitutes the "dark side" of the universe. Some 25% of the total energy density clumps under the influence of gravity, forming the mysterious dark matter whose existence we infer from observations of galaxies and the cosmic microwave background radiation (CMBR). Moreover, 70% of the total energy density appears to be in the form of an even more mysterious dark energy, with negative pressure, which accelerates the expansion of the universe. However, we do not know the nature of this dark matter and we know even less about this dark energy.

We do know that dark matter is not made of baryons (protons and neutrons), because the baryon density, inferred from the primordial abundance of light elements or the CMBR, is much lower than the total matter density. This conclusion has been confirmed by unsuccessful attempts to observe dark baryons, such as those in the form of planetary-sized massive compact halo objects (MACHOs). Light massive neutrinos are also ruled out: Cosmology constrains the sum of their masses to be less than half an electron volt (*1*, *2*). Most evidence now points to "cold" dark matter (i.e., particles that are nonrelativistic at the time of galaxy formation). The combination of cold dark

University of California, Berkeley, CA 94720, USA. E-mail: sadoulet@cosmology.berkeley.edu

matter with inflation and dark energy leads to an impressive account of how structure formed in the universe, in accordance with most observations.

From the point of view of particle physics and gravity, dark matter may be the most compelling evidence that there is physics "beyond the standard model." We know, for instance, that despite its success in describing strong interactions, quantum chromodynamics violates charge-parity symmetry, in strong disagreement with experiment. The most elegant method to deal with this flaw leads to the prediction of a new particle, the axion, that may explain the dark matter. Axions from the halo of our Galaxy can be detected by scattering off the virtual photons of a magnetic field. They would produce real photons in the frequency range 100 MHz to 100 GHz, depending on the axion mass. By arranging for these photons to excite a resonance in a finely tuned radio-frequency cavity kept at low temperature, we can bring the expected feeble signal above the noise level of the best modern amplifiers. With this method, the ADMX (Axion Dark Matter Experiment) group is reaching cosmologically interesting limits (*3*) in the mass range of 2 to 2.2 $\mu eV/c^2$ (where c is the speed of light) for one generic class ("KSVZ") of axion models. During the next 2 years, the group plans to explore the 10^{-6} to 10^{-5} eV/c^2 interval and then increase the sensitivity of its experiment to include other axion types (e.g., "DFSZ") with superconducting quantum interference device–based radio frequency amplifiers. However, higher mass ranges (up to 10^{-3} eV) may be difficult to probe with this method.

Weakly interactive massive particles (WIMPs) form another candidate class. They arise naturally if we assume that the cold dark matter is made of massive particles that were once in thermal equilibrium with ordinary matter in the early universe. As the universe expanded and cooled, they might have dropped out of equilibrium when they were non-relativistic. In this case their present density would be inversely proportional to their interaction rate. To explain dark matter, we need interaction rates typical of the electroweak scale, hence their name. Inversely, we know that the very successful electroweak unification is unstable. To stabilize it, it is necessary to introduce new physics at the same scale, which could be supersymmetry, compact or warped additional dimensions, or the "little Higgs." The least massive of the new particles introduced is usually stable and interacts at the level needed to form dark matter. For very different reasons, particle physics and cosmology then lead to the same concept of WIMPs.

It is therefore attractive to detect WIMPs from the halo of our Galaxy, for example, by elastic scattering on a suitable target. The experimental challenges appear daunting. Expected rates are on the order of one event per kilogram of target per week or per month, much smaller than radioactivity rates in the purest materials. Moreover, the typical recoil energy of 15 keV of scattering on nuclei is very small. Extremely sensitive detector technologies are therefore required with the ability to actively reject the radioactive backgrounds. If the ambient fast neutron level can be sufficiently reduced (by moderation by hydrogenated material and a location deep enough underground), WIMPs are the only known potential source of nuclear recoil events. We can then use unique characteristics of nuclear recoils in terms of pulse shape of the signal or ratios of ionization, scintillation, or total energy deposition to reject electron recoils due to gamma-ray or electron interactions. The ultimate goal is to construct a background-free experiment where this rejection occurs event by event.

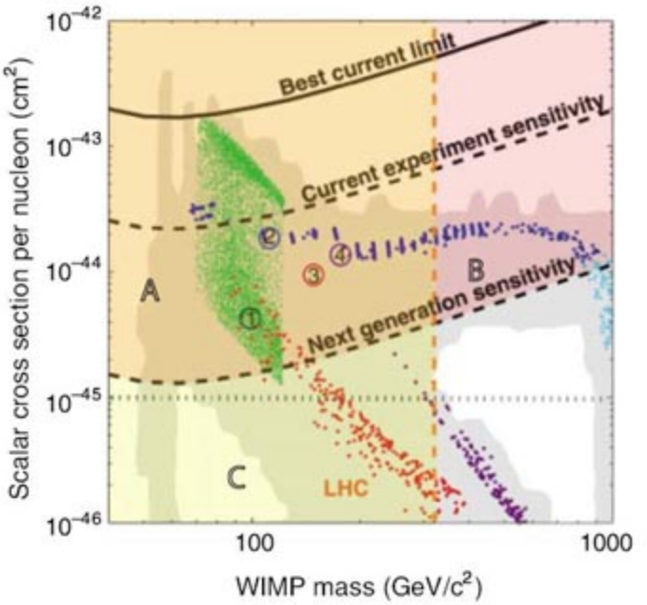


Fig. 1. Sensitivity of current and future experiments to scalar WIMP interactions (*7*). The upper curve is the present best limit (CDMS 2005). The dashed curves show expected sensitivities of the current generation of experiments and of the next generation. The shaded region is the general region of supersymmetry (*8*) compatible with observations. The benchmarks of (*6*) are identified by number. Colored points denote generic regions of supersymmetric models (green, bulk region; blue and cyan, focus point region; purple, co-annihilation; red, funnel). The Large Hadron Collider (LHC) will have difficulty reaching WIMP masses above 350 GeV.

How would we know that we are observing WIMPs? In addition to being nuclear recoils, WIMP interactions should be single scatters uniformly distributed throughout the detector volume. In contrast, neutrons and gamma rays scatter every few centimeters for dense detectors. Eventually, linkage to the galaxy can be established through recoil directionality (which requires low-density media because the recoil length is about 10^{-5} g/cm^2, or a few hundred angstroms in germanium) and yearly modulation of the signal (a few-percent effect that requires several thousands of events to establish). Here, we review the techniques proposed to achieve these goals [see also (*4*, *5*)].

A first generation of experiments attempted to use germanium at 77 K and sodium iodide (NaI), but their sensitivity was limited by the lack of active background rejection. The DAMA (Dark Matter) group claims to have observed WIMPs on NaI through annual modulation. However, their assertion suffers from internal inconsistencies and has not been independently confirmed.

Phonon-mediated (sometimes called "cryogenic") detectors played a pioneering role in demonstrating nearly background-free performance with target masses of around 1 kg. The EDELWEISS (Experience pour Detecter les WIMPs en Site Souterrain) and CRESST (Cryogenic Rare Event Search with Superconducting Thermometers) experiments sense phonons produced in the interactions in the form of a temperature increase measured by highly sensitive thermistors, whereas the CDMS (Cryogenic Dark Matter Search) experiments detect the athermal phonons through their breaking of Cooper pairs in superconducting films covering the surface. These experiments operate at temperatures of 10 to 40 mK so that these athermal phonons are not overwhelmed by thermally excited ones. In addition to their high energy sensitivity, phonon-mediated detectors can discriminate against electron recoils by combining their phonon measurement with ionization measurement at low electric field or with scintillation. Because of the additional information provided by athermal phonons and the additional rejection they provide against surface events, the CDMS II experiment is roughly a factor of 10 more sensitive than any other in the world and begins to enter into the cross section expected for supersymmetry (Fig. 1). In combination with the limits from high-energy solar neutrinos, these results increase the tension with the DAMA claim. Phonon-mediated experiments now in operation use target masses of around 5 kg, with sensitivities 10 times the present limit. New proposals such as SuperCDMS are aiming for 25-kg targets, which should be able to reach 10^{-45} cm^2 per nucleon for a WIMP mass of 60 GeV/c^2, in the next 6 years. The challenge will be to extrapolate these methods to the 1000-kg scale.

Substantial progress has also been made with the use of noble liquids (Ne, Ar, Xe). By comparing scintillation and ionization, it is possible to distinguish nuclear and electron recoils. This requires the amplification of the ionization signal, e.g., by extracting the electrons from the liquid into the gas and having them scintillate in a high–

electric field region. This amplification has now been achieved in three different experiments—ZEPLIN II (Zoned Proportional Scintillation in Liquid Noble Gases) and XENON 10 in Xe, and WARP (WIMP Argon Program) in Ar. Another breakthrough recently occurred when it was realized that for Ne and Ar the pulse shape of nuclear recoils drastically differs from that of an electron recoil. It is then possible to consider a scintillation-only scheme: A sphere is filled with a noble liquid, and the sphere's inner surface is covered with a high density of photomultipliers. A fiducial region protected from external radioactivity can then be defined in the center of the detector, and pulse shape discrimination allows the recognition of nuclear recoils. This scheme is used by the CLEAN (Cryogenic Low-Energy Astrophysics with Neon, which may initially use argon) and XMASS (Xenon Neutrino Mass Detector) proposals. These techniques may gracefully scale to high target masses. However, it remains to be seen what thresholds can be reached.

Gaseous detectors are another option. At low pressure, or higher pressure with sufficiently dense pixels, it should be possible to detect the direction of the recoil. Typically these devices drift charges over long drift length in a time projection chamber. Carrier diffusion must be limited, for example, by a negative ion drift technique. The DRIFT (Directional Recoil Identification from Tracks) group is currently testing underground a cubic-meter prototype at 40 torr (167 g of CS_2). But it is clear that to reach the hundreds of kilograms that are needed, very large chambers of several thousand cubic meters will be needed with some 10^9 pixels, 1 mm on the side.

Metastable systems may enable the construction of detectors that are sensitive to the high energy density created by nuclear recoils and not to photons. One may use Freon droplets [PICASSO (Project in Canada to Search for Supersymmetric Objects), SIMPLE (Superheated Instrument for Massive Particle Experiments)] or a very stable bubble chamber [COUPP (Chicagoland Observatory for Underground Particle Physics)]. Unfortunately, such detectors are also sensitive to alpha interactions and any nucleation agent such as dust. With sufficient purity, it should be possible to produce inexpensive large-mass detectors that are sensitive only to WIMPs. Although these detectors may reach interesting upper limits quickly, they may lack the redundancy needed to substantiate a signal.

These searches complement other experiments, in particular at the Large Hadron Collider (LHC) scheduled to start operation in 2008. Note that the detection of missing energy events at the LHC would not be a sufficient proof that the nature of dark matter has been deciphered. Particles produced at the LHC might be unstable and decay into superWIMPs (such as gravitinos) impossible to detect directly. The simultaneous detection of WIMPs in dark matter experiments and at the LHC (Fig. 1) would open a very rich field of investigation (*6*), but there are scenarios where WIMPs are accessible to only one technique. On a similar time scale, the GLAST (Gamma-ray Large Area Space Telescope) satellite may detect gamma-ray emission from WIMP annihilation in galaxies (possibly two to five identifications within 5 years). Emission from WIMP annihilation is also expected from the galactic center, but the interpretation of a signal may not be unique.

Deciphering the nature of the dark matter in the universe is important not only for astrophysics and cosmology but also for particle physics and gravity. The most unambiguous results may come from the direct detection of halo dark matter in the laboratory. For axions and WIMPs, current technologies are reaching sensitivity levels of cosmological interest and a number of novel detection schemes are in development. The roadmap for WIMP searches appears to be clear: It is important to increase the target mass rapidly while maintaining zero background. This could be achieved by a combination of searches based on the demonstrated phonon-mediated sensors and on promising new technologies such as noble liquids, with several experiments cross-checking each other. The region of 10^{-45} cm^2 per nucleon appears particularly interesting and reachable within the next years, by which time the LHC will test supersymmetry or additional dimensions, and GLAST will fly. If we are lucky, we may indeed be at the brink of a discovery. If such a discovery occurs, by linking the recoil directions to the galaxy, we may even confirm that the observed events are indeed due to dark matter.

References and Notes

1. U. Seljak *et al.*, *Phys. Rev. D* **71**, 103515 (2005).
2. M. S. Turner, *Science* **315**, 59 (2007).
3. L. D. Duffy *et al.*, *Phys. Rev. D* **74**, 012006 (2006).
4. R. J. Gaitskell, *Annu. Rev. Nucl. Part. Sci.* **54**, 315 (2004).
5. D. Akerib *et al.*, http://arxiv.org/abs/astro-ph/0605719 (2006).
6. E. A. Baltz, M. Battaglia, M. E. Peskin, T. Wizanski, http://arxiv.org/abs/hep-ph/0602187 (2006).
7. J. Filippini, personal communication.
8. E. A. Baltz, P. Gondolo, *J. High Energy Phys.*, JHEP10(2004)052 (2004).

9. Supported by NSF grant PHY 05-04224. The author is a member of CDMS, which is funded by NSF and the U.S. Department of Energy.

10.1126/science.1136639

PERSPECTIVE

Neutrino Astrophysics: A New Tool for Exploring the Universe

Eli Waxman

In the past four decades a new type of astronomy has emerged, where instead of looking up into the sky, "telescopes" are buried miles underground or deep under water or ice and search not for photons (that is, light), but rather for particles called neutrinos. Neutrinos are nearly massless particles that interact very weakly with matter. The detection of neutrinos emitted by the Sun and by a nearby supernova provided direct tests of the theory of stellar evolution and led to modifications of the standard model describing the properties of elementary particles. At present, several very large neutrino detectors are being constructed, aiming at the detection of the most powerful sources of energy and particles in the universe. The hope is that the detection of neutrinos from these sources, which are extra-Galactic and are most likely powered by mass accretion onto black holes, will not only allow study of the sources, but, much like solar neutrinos, will also provide new information about fundamental properties of matter.

Neutrino astronomy was initiated as an attempt to provide a direct experimental test for the theory of stellar evolution (*1*). According to this theory, the Sun is powered by the nuclear fusion of hydrogen into helium, which takes place deep in the solar interior. The mass of four H atoms is larger than that of the He atom into which they fuse. The excess mass m is converted to energy, according to $E = mc^2$ (c is the speed of light), which keeps the Sun shining. It was suggested in the mid-1960s that one could test this model by searching for neutrinos, which were predicted to be emitted by the fusion process. Unlike photons that are emitted from the Sun's surface, the weak interaction of neutrinos with matter allows them to escape from the Sun's core and directly reach detectors on Earth.

The weak interaction of neutrinos with matter also implies that they are very difficult to detect, requiring the construction of detectors with several kilotons of detecting medium. Although the probability that a neutrino passing through kilotons of matter would interact within the detector, or be "captured," is very small, the large flux of neutrinos from the Sun, some 100 billion neutrinos per square centimeter per second, allows hundreds of them to be detected every year. In addition to being very massive, all detectors are also buried deep underground. At the surface of Earth there is a large flux of high-

Physics Faculty, Weizmann Institute of Science, Rehovot 76100, Israel. E-mail: waxman@wicc.weizmann.ac.il

energy particles. Such particles are produced mainly by the interaction of cosmic-rays, high-energy particles produced in space, with the atmosphere. Penetration of high-energy particles into the detector may lead to interactions that would mimic neutrino interactions. Burial of the detector deep underground suppresses this background, because only neutrinos can penetrate deep enough into Earth to reach the detector.

The detection of solar neutrinos was an impressive confirmation of the hypothesis of a nuclear fusion origin of stellar energy. However, it also posed a challenge: The measured neutrino flux was roughly one-half that predicted by theory. Shortly after this discrepancy was first reported in 1968, it was suggested to be due to shortcomings of the standard model describing the properties of elementary particles (*2*). Neutrinos come in three types, or "flavors": electron-type (ν_e), muon-type (ν_μ), and tau-type (ν_τ). It was proposed that about half of the neutrinos, which are produced in the Sun by nuclear fusion and are all of electron-type, change their flavor to ν_μ or ν_τ as they propagate to Earth. Such flavor conversion, commonly termed "oscillation," was not expected according to the standard model and would explain why neutrino detectors sensitive to ν_e only would miss about half of the solar neutrino flux.

The oscillation explanation was confirmed in 2001 with the detection of the "missing" ν_e flux in the form of ν_μ and ν_τ flux by an experiment sensitive to all flavors (*3*). Independent evidence for neutrino oscillations came from measurements of atmospheric neutrinos, produced by cosmic-ray interactions in the atmosphere, which indicate conversion of ν_μ to ν_τ (*3*). Neutrino oscillations are the first, and so far only, experimental phenomenon not accounted for by the standard model. It is most naturally explained by a model in which three neutrinos with different masses exist—say, ν_1, ν_2, and ν_3 with masses m_1, m_2, and m_3, respectively—and in which neutrinos of different flavors are in fact composed of different mixtures of ν_1, ν_2, and ν_3. ν_e, for example, is a roughly equal mixture of ν_1 and ν_2 with little (if any) contribution of ν_3.

After the discovery of neutrino oscillations by observing natural (solar and atmospheric) neutrino sources, oscillations were also measured with neutrinos produced in nuclear reactors and particle accelerators. Oscillation measurements provide constraints on the neutrino "mixing parameters" (*2*, *3*), that is, on the composition (in terms of $\nu_{1,2,3}$) of neutrinos of different flavors, and on the mass-squared differences, $m_2^2 - m_1^2 = 8 \times 10^{-5}$ (eV/c^2)^2 and $|m_3^2 - m_2^2| = 2 \times 10^{-3}$ (eV/c^2)^2. Here, masses are given in energy units, where $m = E/c^2$; 1 eV is the typical binding energy of molecules and corresponds to roughly one-millionth of the electron mass, $m_e c^2 = 0.5 \times 10^6$ eV. Oscillation experiments cannot determine the absolute values of the masses, and current data do not allow one to discriminate between the two "hierarchies," $m_1 < m_2 < m_3$ and $m_3 < m_1 < m_2$. An upper limit on the mass of the most massive neutrino, $m \leq 2$ eV/c^2, is set by measurements of radioactive decay of tritium (*4*). A similar upper limit is obtained from surveys of the large-scale distribution of galaxies: The universe is filled with a "neutrino background," a relic of the big bang, and if neutrinos were too massive, they would have suppressed the formation of large-scale structures in the universe (*5*).

A model explaining the origin of neutrino masses and mixing does not yet exist (*3*). To construct such a model, large radioactive-decay experiments are planned in order to measure the absolute neutrino mass scale (*4*), and large oscillation experiments involving reactors and specially designed accelerator beams are planned for determining the mass hierarchy (and for accurate determination of the mixing parameters) (*6*). These experiments will also try to ascertain whether the mixing properties of neutrinos and of their antiparticles, antineutrinos, are identical. Answering these questions would be important for the construction of a model accounting for neutrino mass and mixing. It may also be relevant for answering another open question—why our universe appears to be composed mainly of particles and not of antiparticles (*7*).

According to the theory of stellar evolution, stars more massive than the Sun by a factor of 10 or more end their lives with an explosion, a supernova, that ejects most of the star's mass and leaves behind a dense "neutron star" remnant of roughly 1 solar mass. Theory predicts that most of the energy generated by the explosion would be carried away from the star by neutrinos. This prediction was confirmed (*1*) with the detection in 1987 of neutrinos emitted by the supernova SN 1987A, which exploded in the Large

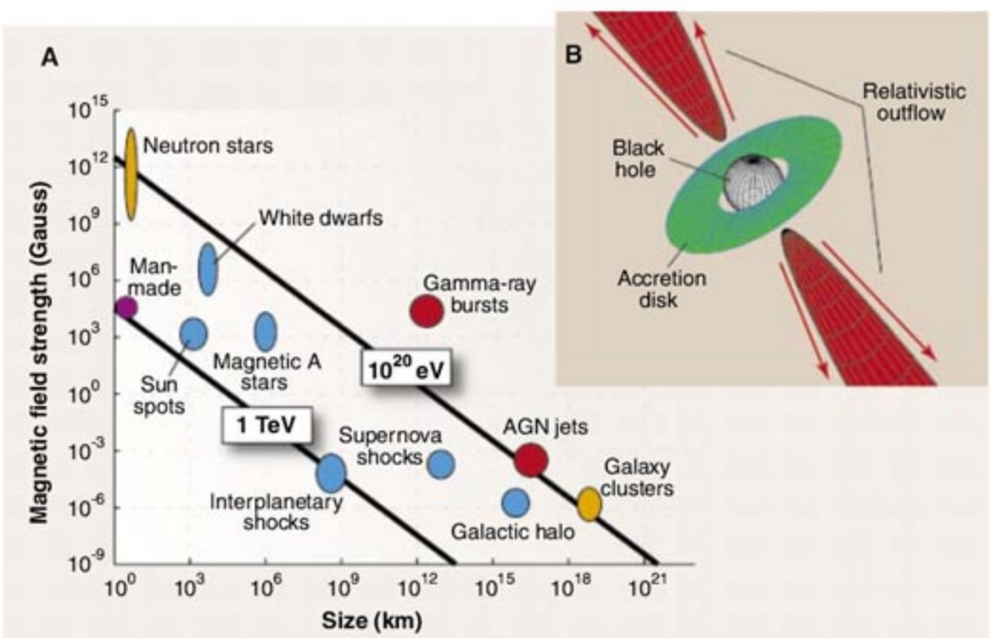


Fig. 1. (**A**) Charged particles are confined to their astrophysical accelerators by magnetic fields. Magnetic confinement requires the product of field strength and accelerator size to exceed a value, which increases with particle energy. The figure shows the size and magnetic field strength of possible sites of particle acceleration. (The magnetic field is measured in Gauss units, where the Earth's magnetic field is ~1 G.) Proton acceleration to 1 TeV or 10^{20} eV is possible only for sources lying above the appropriately marked lines. This is a necessary, but not sufficient, requirement: Proton acceleration to 10^{20} eV is impossible in galaxy clusters (because the acceleration time in these objects is larger than the age of the universe) and unlikely in highly magnetized neutron stars (due to severe energy losses). The characteristics of terrestrial man-made accelerators, which are planned to reach ~ 1 TeV, are shown for comparison. (**B**) GRBs and AGN are believed to be powered by black holes. The accretion of mass onto the black hole, through an accretion disk, releases large amounts of gravitational energy. If the black hole is rotating rapidly, the rotational energy may also be released by slowing the black hole through interaction with the disk. The energy released drives a jetlike relativistic outflow. The observed radiation is produced as part of the energy carried by the jets is converted, at a large distance from the central black hole, to electromagnetic radiation. GRBs are believed to be powered by ~1–solar mass black holes with jets extending to distances larger than the size of the solar system, producing short (typically 1- to 100-s-long) flashes of luminosity exceeding that of the Sun by 19 orders of magnitude. AGN are powered by million– to billion–solar mass black holes residing at the centers of distant galaxies, with jets extending to distances larger than the size of a galaxy, producing a steady luminosity exceeding that of the Sun by 12 orders of magnitude.

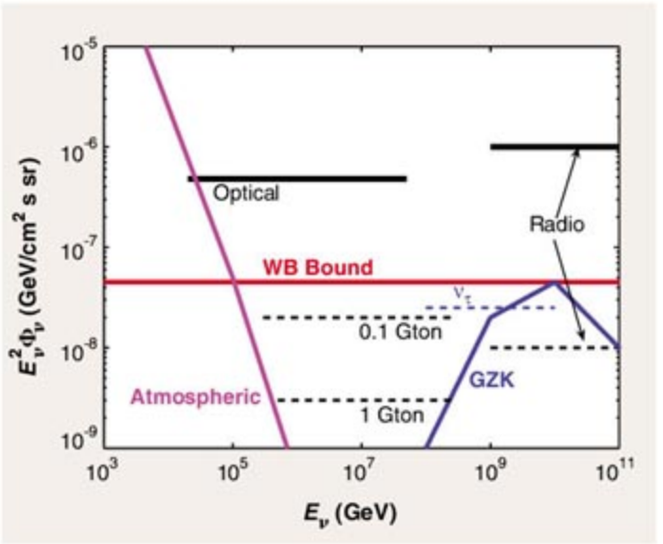


Fig. 2. The cosmic-ray upper bound on the extra-Galactic high-energy neutrino intensity ($\nu_\mu + \nu_\tau$ assuming ν_e:ν_μ:ν_τ = 1:1:1 flux ratios), compared with experimental upper bounds (solid lines) provided by optical detectors under water or ice and by radio detectors, and with the expected sensitivity (dashed lines) of various detectors: 0.1 gigaton and 1 gigaton under water/ice optical detectors, radio detectors, and ground arrays of particle detectors (sensitivity to ν_τ). The intensity Φ_ν is the number of neutrinos of given energy E_ν (measured in GeV = 1000 MeV) crossing in 1 s a unit area (1 cm^2) of a detector observing a solid angle of 1 sr of the sky. A detailed description of the experiments is given in (*8*). The curve marked GZK shows the neutrino intensity expected to be produced by the interaction of high-energy cosmic-ray protons with the cosmic microwave background, the relic radiation of the big bang. Also shown is the atmospheric neutrino intensity, which is produced by cosmic-ray interactions in the atmosphere and constitutes the main background.

Magellanic Cloud, a small satellite galaxy of our own Galaxy, the Milky Way, lying at a distance of some 150 thousand light-years away.

The characteristic energy of neutrinos produced in the Sun or in supernova explosions is on the order of megaelectron volts (1 MeV = 10^6 eV), which is the characteristic energy released in the fusion or fission of atomic nuclei. The detection of MeV neutrinos from sources well outside our local Galactic neighborhood, at distances ranging from several million light-years (the typical distance between galaxies) to several billion light-years (the size of the observable universe), is impossible with present techniques. To extend the distance accessible to neutrino astronomy to the edge of the observable universe, several high-energy neutrino telescopes are currently being constructed deep under ice or water. These telescopes are designed for the detection of neutrinos with energies exceeding terraelectron volts (1 TeV = 10^{12} eV) and are planned to reach effective masses exceeding 1 gigaton (*8*).

The sources targeted by high-energy (≥1 TeV) neutrino detectors are "cosmic accelerators," in which particles are accelerated to extreme energies. The existence of cosmic-rays, high-energy particles that are produced in astrophysical objects and are observed as they hit and interact with Earth's atmosphere, has been mentioned above. The sources of these particles have not yet been identified, and the mechanisms that lead to particle acceleration are not well understood. One of the major goals of ≥1-TeV neutrino detectors is to resolve these open questions.

Particle-acceleration theories are most challenged by the highest-energy particles observed (*9*). These particles are most likely protons, and their energy exceeds 10^{20} eV, or 100 million TeV. Although there are a variety of astrophysical objects suspected of being "cosmic accelerators" (Fig. 1A), only two types of sources are known that may be capable of accelerating protons to 10^{20} eV: gamma-ray bursts (GRBs) and active galactic nuclei (AGN). These objects lie at cosmological distances, billions of light years away and are the brightest known objects in the universe (Fig. 1B). Although GRB and AGN models are generally successful in explaining most observations, they are largely phenomenological, and major questions remain open. These include the mechanisms by which gravitational energy is harnessed to power the sources, and the mechanism of particle acceleration.

A direct association of cosmic-rays with their sources is difficult: Magnetic fields in interstellar and intergalactic space deflect the electrically charged cosmic-rays, which, therefore, do not travel on straight lines and do not point back to their sources. Neutrinos, on the other hand, are electrically neutral and therefore travel on straight lines and do point back to their sources. Whatever the cosmic accelerators are, they are expected to be sources of high-energy neutrinos and therefore to be identifiable by their neutrino emission. This expectation is based on the fact that the interaction of high-energy cosmic-rays with radiation or matter leads to the production of neutrinos. High-energy protons, for example, may interact with photons to produce pions, particles that decay and produce muon and electron neutrinos.

Observations of high-energy cosmic-rays provide a means for estimating the expected high-energy neutrino flux and hence the detector size required to measure it. The observed cosmic-ray flux sets an upper bound to the neutrino flux produced by extra-Galactic sources (*9*), which implies that gigaton neutrino telescopes are needed to detect the expected extra-Galactic flux in the energy range of ~1 to ~1000 TeV, and much larger effective mass is required at higher energy (Fig. 2). A flux comparable to the bound at ~1 to ~1000 TeV would produce hundreds of events per year in a gigaton detector. A few tens of events per year are expected in a gigaton telescope if GRBs are the sources of high-energy protons. These events will be correlated in time and direction with GRB photons, allowing for an essentially background-free experiment.

Detection of high-energy neutrinos with the next generation of telescopes will probe the most powerful cosmic accelerators, including GRBs and AGN, and will allow study of the physical mechanisms powering them. It will also provide new tests of neutrino oscillation theory and probes of fundamental physics that are not available with terrestrial, man-made sources: Flavor measurements of high-energy neutrinos will contribute to the determination of the mixing parameters (e.g., to resolving the mass hierarchy ambiguity and to testing for differences in particle and antiparticle behavior) (*10*). The angular dependence of neutrino detection rate may allow testing for deviations from standard model predictions of the neutrino-nucleon interaction cross section at energies not accessible to terrestrial accelerators (*8*). Detection of neutrinos from GRBs could be used to test the simultaneity of neutrino and photon arrival to an accuracy of ~1 s. This would allow the validity of the underlying assumption of special relativity—that photons and neutrinos have the same limiting speed—to be determined with an accuracy of one part in 10^{17}, and the validity of the weak equivalence principle—the basic assumption of general relativity according to which photons and neutrinos should experience the same time delay as they pass through a gravitational potential—to be measured with an accuracy better than one part in 10^6 (*9*). Previous applications of these ideas to supernova 1987A yielded much weaker upper limits, on the order of 10^{-8} and 10^{-2}, respectively (*1*). Finally, neutrino telescopes may contribute to the detection of "dark matter," unseen particles that were not detected in laboratories on Earth and are believed to contain most of the mass in the universe (*11*), through the detection of neutrinos produced by annihilation of dark-matter particles.

References

1. J. N. Bahcall, *Neutrino Astrophysics* (Cambridge Univ. Press, New York, 1989).
2. S. M. Bilenky, *Phys. Scr.* **T121**, 17 (2005).
3. M. C. Gonzalez-Garcia, Y. Nir, *Rev. Mod. Phys.* **75**, 345 (2003).
4. C. Weinheimer, *Phys. Scr.* **T121**, 166 (2005).
5. Ø. Elgarøy, O. Lahav, *N. J. Phys.* **7**, 61 (2005).
6. M. Lindner, *Phys. Scr.* **T121**, 78 (2005).
7. T. Yanagida, *Phys. Scr.* **T121**, 137 (2005).
8. F. Halzen, *Science* **315**, 66 (2007).
9. E. Waxman, *Phys. Scr.* **T121**, 147 (2005).
10. W. Winter, *Phys. Rev. D* **74**, 033015 (2006).
11. B. Sadoulet, *Science* **315**, 61 (2007).

10.1126/science.1136638

PERSPECTIVE

Neutrino Astrophysics Experiments Beneath the Sea and Ice

Francis Halzen

Neutrino astronomy beyond the Sun was first imagined in the late 1950s. A neutrino detector at the bottom of Lake Baikal, the deployment of detectors in the Mediterranean Sea, and the construction of a kilometer-scale neutrino telescope at the South Pole exemplify current efforts to realize this dream.

Our universe exhibits nuclear processes far more violent than those that can be created by earthbound particle accelerators. Throughout the cosmos, nature accelerates elementary particles to energies in excess of 10^{20} electron volts, equivalent to a macroscopic energy of 50 joules carried by a single elementary particle. We have no idea where these particles, most likely protons, originate or how they are accelerated to such high energies.

There are several problems when using high-energy particles, or cosmic rays, to carry out astronomy. Because cosmic rays are electrically charged, their paths become scrambled by pervasive galactic and, in some cases, intergalactic magnetic fields, so their arrival directions at Earth do not reveal their exact origin. This is why the cosmic ray puzzle persists almost a century after the discovery of radiation from space. The flux of particles with energies high enough to undergo minimal deflection is so small that sources have proved impossible to observe directly up to now. The Auger detector covering several thousand square kilometers of the high plateau in Argentina may collect such events with sufficient statistics (*1*).

Cosmic rays are also challenging astronomical messengers for another reason: They self-destruct in collisions with universal microwave background photons. As a result, they only reach us from our nearby cosmic neighborhood. Very-high-energy photons share this problem too. For example, greatly improved techniques to collect TeV-energy photons, using the atmosphere as the detector, now probe the universe to redshifts of only $z \sim 0.1$.

However, after the discovery of cosmic neutrinos in the 1950s in the radiation of nuclear reactors, many realized that they did not have the same limitations as charged cosmic particles and photons. Neutrinos had the potential to be ideal cosmic messengers. Unfortunately, building a neutrino telescope has turned out to be a daunting technical challenge (*2*).

With essentially no mass and no electric charge, the neutrino is similar to the photon as a cosmic messenger. It differs in one important attribute, however: Its interactions with matter are extremely feeble. This can be advantageous in that high-energy neutrinos may reach us unscathed from the edge of the universe, from the inner neighborhood of black holes, and, hopefully, from the nuclear furnaces where cosmic rays are born. They may tell us about cosmic sites never "seen" and let us peer into the hearts of black holes.

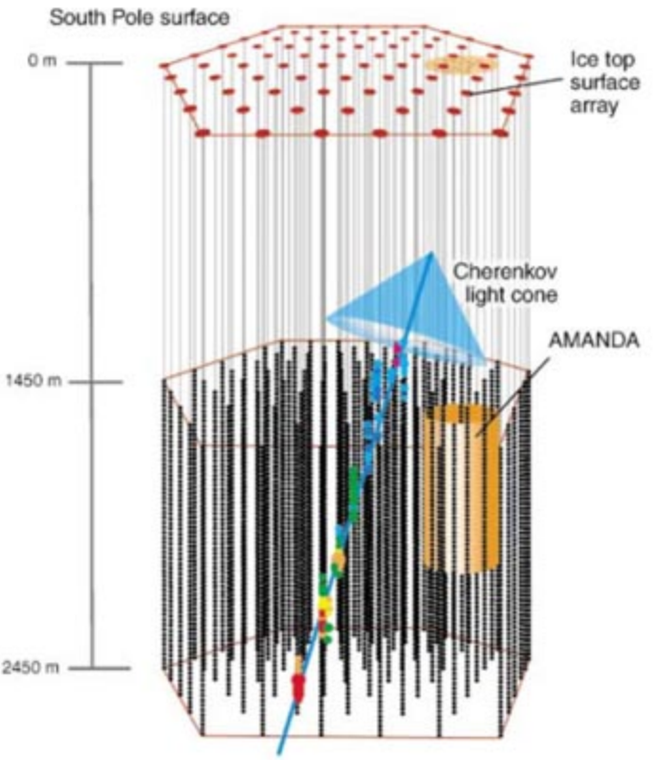


Fig. 1. Design of the IceCube kilometer-scale neutrino observatory. Red dots represent air shower detectors laid out on the surface ice sheet for calibration and shielding of the deep ice detector. IceCube consists of 4800 digital optical sensors (black dots) viewing a cubic kilometer of ice between 1450 and 2450 m. The orange cylinder indicates the volume of ice instrumented by 677 AMANDA sensors. The colored dots show the response of the detector to the Cherenkov light radiated by a simulated 10-TeV muon track. The colored dots (red to purple in rainbow order) indicate the arrival time of the light, red lighting up first; their size is proportional to numbers of photons. The track has been initiated by a neutrino that interacted below the detector after traveling through Earth from a northern source.

Unfortunately, this feeble interaction with matter makes cosmic neutrinos also very difficult to detect. Trillions of neutrinos fly through our body every second. On average, one high-energy neutrino produced in cosmic ray interactions with atmospheric nuclei will stop within each of us in a lifetime. Immense particle detectors are required to collect cosmic neutrinos in sufficient numbers to be statistically significant to pursue science. By the 1970s it was clear that a cubic-kilometer neutrino detector would need to be constructed to reveal the neutrinos produced by cosmic rays interacting with background microwave photons. Up-to-date estimates for observing cosmic sources such as quasars or gamma ray bursts, unfortunately, point at the same exigent requirement (*3*).

Given the size of the detector required, efforts concentrated on transforming large volumes of natural water into Cherenkov detectors that catch the flashes of light produced by the rare neutrinos that interact in or near the detector. After an effort that spanned more than two decades, building the Deep Underwater Muon and Neutrino Detector (DUMAND) in the sea off the main island of Hawaii unfortunately failed (*2*). However, it paved the road for later efforts by developing many of the detector technologies in use today, and by inspiring the deployment of a smaller instrument in Lake Baikal (*4*). Its successful operation bodes well for efforts to commission neutrino telescopes today in the Mediterranean: ANTARES (Astronomy with a Neutrino Telescope and Abyss Environmental Research) (*5*) and NESTOR (Neutrino Extended Submarine Telescope with Oceanographic Research) (*6*).

The first telescope on the scale originally envisaged by the DUMAND collaboration has been realized instead by transforming a large volume of the extremely transparent natural deep Antarctic ice into a particle detector, the Antarctic Muon and Neutrino Detector Array (AMANDA). AMANDA, in operation since 2000, represents a proof of concept for the kilometer-scale neutrino observatory, IceCube, now under construction (*7*).

Even extremely high-energy neutrinos will routinely stream through the detectors without leaving a trace; the unlucky one that makes a direct hit on a nucleus in the water or ice will create muons as well as electromagnetic and hadronic

Department of Physics, University of Wisconsin, Madison, WI 53703, USA. E-mail: francis.halzen@icecube.wisc.edu

secondary particle showers familiar from accelerator experiments. The charged remnants will radiate a glow of blue light, dubbed Cherenkov radiation, that will spread through the natural ice over hundreds of meters. The origin of this radiation is the same as that of the blue glow shining from the water shielding nuclear reactors. Neutrino astronomers embed optical sensors into Antarctic ice to detect the faint light from a nuclear reaction initiated by a single neutrino. The light pattern reveals the direction of the neutrino, making neutrino astronomy possible. Among the secondaries, muons are of special interest because the mean free path of the most energetic muons can exceed 10 km. The effective detector volume thus exceeds the instrumented volume for muon neutrinos.

In general, a neutrino telescope must be (i) kilometer-size to detect the low fluxes of neutrinos from cosmic sources, (ii) transparent enough to allow light to travel through a widely spaced array of optical sensors, (iii) deep enough to be shielded from surface light and radiation, and (iv) affordable. Only deep dark oceans and glaciers of ice satisfy these constraints. Pure, highly transparent, and free of radioactivity, Antarctic polar ice has turned out to be an ideal medium to detect neutrinos. The difficulty of the remote site has been overcome by exploiting the infrastructure of the U.S. National Science Foundation's Amundsen-Scott South Pole Station.

AMANDA is the initial stage and proof of concept for a kilometer-scale neutrino observatory, IceCube, now under construction at the South Pole. IceCube will instrument a cubic kilometer of ice surrounding the AMANDA detector (Fig. 1). Its basic detector component is a photomultiplier housed in a glass pressure vessel, somewhat larger than the size of a basketball (Fig. 2B). Photomultipliers transform the Cherenkov light from neutrino interactions into electric signals by the photoelectric effect. The signals are captured by a computer chip that digitizes the shape of the current pulses and sends the information to the computers collecting the data, first by cable to the "counting house" at the surface of the ice sheet and then via magnetic tape or, in the case of more interesting events, by satellite to the IceCube Data Warehouse in Madison, Wisconsin. One can think of IceCube as 4800 freely running computers sending time-stamped digitized images of the light they detect to the surface. The local clocks in the sensors are kept calibrated with nanosecond precision. This information allows the scientists to reconstruct neutrino events and infer their arrival directions and energies. The detector components transform a cubic kilometer of ice at a depth of 1450 to 2450 m into a cosmic neutrino detector (i.e., 1 mile below the surface and ¼ mile above bedrock).

Optical sensors produced at collaborating institutions in the Northern Hemisphere are shipped to the international Antarctic center in Christchurch, New Zealand. These are later transported to the South Pole by way of the port at McMurdo, Antarctica. Drillers use a 5-MW high-pressure jet of hot water to melt a hole in the ice, roughly half a meter wide and 2.5 km deep, in less than 2 days. Because ice is an excellent insulator, the water does not freeze for several days, ample time to deploy the optical sensors attached to cables that will power them and will also transmit their digital signals to the surface (Fig. 2B). Each of 80 holes will hold 60 sensors evenly spread over 1 km between depths of 1450 and 2450 m.

With some 650 optical sensors in place since February 2000, the existing AMANDA detector has been collecting neutrinos at a steady rate of four per day. These "atmospheric neutrinos" are the by-product of collisions of cosmic rays with the nitrogen and oxygen in the northern atmosphere. Note that at the South Pole one observes neutrinos that originate in the Northern Hemisphere, looking through Earth (used as a filter) to select neutrinos from other particles. No photons, or any other particles besides neutrinos, can traverse the whole planet to reach the detector. The signals from the atmospheric neutrinos do not yet yield information about astronomy, but they are calculable and can be used to prove that the detector performs as expected. As in conventional astronomy, AMANDA will have to look beyond the atmosphere for cosmic signals; AMANDA data are now scrutinized for hot spots in the northern sky that may signal cosmic sources.

Starting in the Antarctic summer of 2004–2005, IceCube deployments have been steadily augmenting the AMANDA instrumentation. As of January 2006, IceCube consists of 604 digital optical modules distributed over nine strings and 32 surface cosmic ray detectors. The hardware and software worked "out of the box" and revealed the first atmospheric neutrinos in early February 2006. The collaboration is now analyzing its first 6 months of data. Over the next four seasons, IceCube will transform the ice into the kilometer-scale neutrino observatory that is required for neutrino astronomy. However, detector elements deliver information as soon as they are deployed, and thus IceCube will deliver a kilometer-square year of integrated observations of the Northern Hemisphere by 2008–2009.

After extensive research and development (R&D) efforts by both the ANTARES and NESTOR collaborations in the Mediterranean, there is optimism that the technological challenges to build neutrino telescopes in deep seawater have now been met. Both Mediterranean collaborations have demonstrated their capability to deploy and retrieve optical sensors. The initial deployments targeted R&D of the detector components and in situ study of the water. The deployed optical sensors could also be operated as a particle detector. Both collaborations have reconstructed downgoing cosmic ray muons with the optical modules that were deployed for R&D tests. Although the instrumentation was too limited to detect neutrinos, both collaborations validated their detector designs by detecting cosmic ray muons.

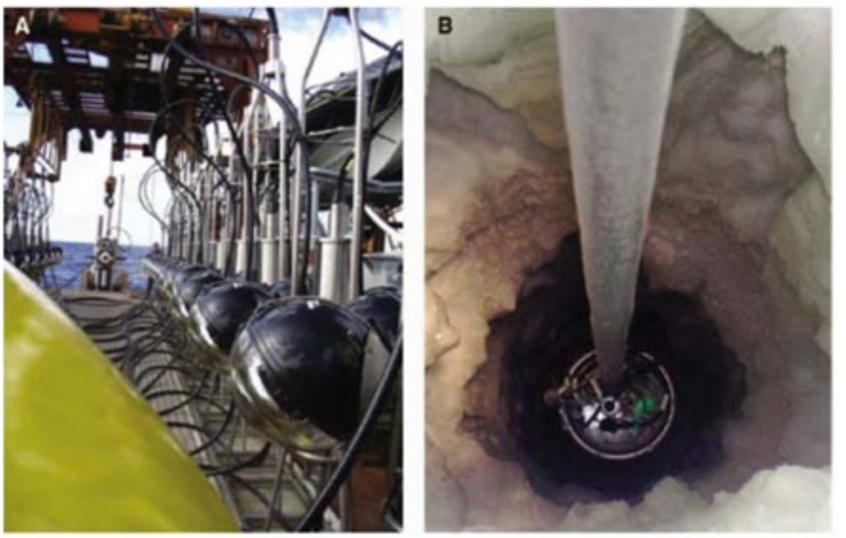


Fig. 2. Deployment of optical sensors (10-inch photomultiplier tubes encased in a centimeter-thick glass pressure housing) by ANTARES in water (**A**) or by IceCube in (temporarily melted) ice (**B**).

The final construction of the ANTARES detector, which will have a similar size as AMANDA, started in February 2006. It is conveniently located at a depth of 2400 m close to the shore near Toulon, France. The detector will consist of 12 strings, each equipped with 75 optical sensors mounted in 25 triplets. The collaboration has by now deployed two strings that were connected by submarine to cables transmitting the data via a junction box to shore. The strings have been successfully and reliably taking data. The completion of the detector is foreseen about 1 year from now. The beginning of operation of ANTARES marks a historic milestone by opening the Southern Hemisphere, and hence the galactic center, for neutrino astronomy. Also, NESTOR is expected to augment its prototype installation in the near future.

Furthermore, a European Union–funded design study dubbed KM3NeT is intended to create a technical design report for the construction of a kilometer-scale detector in the Mediterranean Sea, complementary to IceCube at the South Pole. KM3NeT is a common effort of the Mediterranean

projects, including the Neutrino Mediterranean Observatory (NEMO) in Catania, Italy, that has already done R&D toward a kilometer-scale detector. The 3-year study started early this year. The recent project's inclusion in the Road Map of the European Strategy Forum and Research Infrastructures (ESFRI) represents an important step toward the realization of the project. The start of construction of KM3NeT is envisaged for the beginning of the next decade, in time for concurrent operation with IceCube.

As is the case for conventional telescopes, neutrino telescopes inevitably view the universe through Earth's atmosphere. Cosmic rays interacting with atmospheric nuclei produce a uniform background of neutrinos that must be separated from those of cosmic origin. AMANDA, while too small to reveal cosmic sources, has successfully exploited atmospheric neutrinos as a calibration beam. Whereas AMANDA collected some 5000 neutrinos with energy up to 100 TeV, in only a few years IceCube will collect several hundred thousand neutrino atmospheric events with energies of 0.1 to 1000 TeV, well in excess of those observed in particle physics laboratories. Exploiting this guaranteed beam, neutrino "telescopes" will thus also open a new chapter in particle physics. Particle discoveries with natural neutrino beams include neutrino mass. Even in the absence of a discovery, the experiment will be able to test basic principles such as Lorentz symmetry and the equivalence principle with a sensitivity improved by more than two orders of magnitude over present experiments. They may reveal Planck-scale physics (*8*).

References

1. J. W. Cronin *et al.*, *Sci. Am.* **276**, 32 (January 1997).
2. A. Roberts, *Rev. Mod. Phys.* **64**, 259 (1992).
3. F. Halzen, D. Hooper, *Rep. Prog. Phys.* **65**, 1025 (2002).
4. V. A. Balkanov *et al.*, *Nucl. Phys. Proc.* **118** (suppl.), 363 (2003).
5. J. A. Aguilar *et al.*, *Astropart. Phys.* **26**, 314 (2006).
6. G. Aggouras *et al.*, *Astropart. Phys.* **23**, 377 (2005).
7. A. Achterberg *et al.*, *Astropart. Phys.* **26**, 155 (2006).
8. U. Jacob, T. Piran, http://arxiv.org/abs/hep-ph/0607145 and references therein.

10.1126/science.1136504

PERSPECTIVE

Cosmic Rays: The Highest-Energy Messengers

Angela V. Olinto

The origin of the most energetic particles ever observed, cosmic rays, will begin to be revealed in the next few years. Newly constructed ultrahigh-energy cosmic ray observatories together with high-energy gamma-ray and neutrino observatories are well positioned to unveil this mystery before the centenary of their discovery in 2012. Cosmic ray sources are likely to involve the most energetic phenomena ever witnessed in the universe.

Cosmic rays have a long history, starting in 1912 when Victor Hess lifted electroscopes in balloons to 5-km altitudes and determined that the mysterious ionizing radiation was coming from space and not from Earth. Early cosmic ray physicists used this natural flux of high-energy protons to discover a number of elementary particles, such as the positron, the muon, and the pion, by observing them in cloud chambers and photographic emulsions at high altitudes, where the flux at high energies is less attenuated. By 1938, Pierre Auger showed that very-high-energy cosmic rays trigger extensive air showers in Earth's atmosphere, distributing the original cosmic ray energy among billions of lower-energy particles that arrive together on the ground. In 1962, the Volcano Ranch array led by John Linsley observed a cosmic ray event with an energy of tens of joules or about 10^{20} eV. Four years later, Greisen in the United States (*1*) and Zatsepin and Kuzmin in the USSR (*2*) predicted the abrupt steepening of the cosmic ray spectrum above 10^{20} eV as a result of cosmic ray interactions with the newly discovered cosmic microwave background (CMB). In his landmark article, Greisen announced that the measurement of such a flux steepening would clarify the origin of ultrahigh-energy cosmic rays (UHECRs) by showing its "cosmologically meaningful termination."

A range of different techniques have allowed the observation of cosmic rays from energies just below 10^9 to 10^{20} eV (*3*). Up to 10^{14} eV, direct detection is feasible with balloon and space experiments. Above this energy, the flux is too low for space-based detectors, and cosmic rays are studied by observing their air-shower development. Direct detection shows that at low energies the cosmic ray flux is modulated by the solar cycle through the magnetic heliosphere, which shields the solar system from charged particles below about 10^9 eV. From a few GeV (1 GeV = 10^9 eV) to a few PeV (1 PeV = 10^{15} eV), the cosmic ray spectrum is well described by a power law of spectral index −2.7—i.e., the number of cosmic rays arriving on Earth per unit time, area, solid angle, and kinetic energy, E, is $J(E) \propto E^{-2.7}$. At higher energies, the spectrum steepens to $J(E) \propto E^{-3}$ and the transition region is called the "knee." At about 10^{18} eV the spectrum hardens again, giving rise to a feature named the "ankle." Finally, at about 10^{20} eV, the "cosmologically meaningful termination" predicted by Greisen, Zatsepin, and Kuzmin is expected as these UHECRs lose energy through pion production in interactions with the CMB radiation. This final feature is named the GZK cut-off after its originators. The exact position and shape of each of these features is presently under intense research, because they give clues to the cosmic ray production and propagation mechanisms.

Composition studies at low energies exposed the diffusive history of cosmic ray nuclei as they propagate through the Galaxy. Spallation products of abundant nuclei are much more abundant in cosmic rays than in solar system material; for example, cosmic rays Li, Be, and B—produced mainly by the spallation of C and O—are 5 orders of magnitude more abundant than their solar values. The overabundance, together with spallation cross-sections, shows that cosmic rays have traversed from 5 to 10 g/cm^2 as they propagate in the Galaxy, corresponding to trajectories of ~1 Mpc (equal to 3×10^{24} cm) in length, which is much larger than the thickness of the galactic disk (~ 0.4 pc).

While at energies below 10^{15}eV, cosmic rays are dominated by light nuclei (protons and helium); above the knee, the composition seems to become heavier. This transition to heavier elements is expected in models where cosmic rays propagate diffusively in the galactic magnetic field with a probability of escape that depends on rigidity (i.e., ratio of energy to the charge). Within this picture, the knee would represent the transition from confined trajectories to trajectories that escape the Galaxy and thus produce the change in the spectral index. Tests of this model and alternative proposals are currently under scrutiny by a number of observatories. Leading this effort is the Karlsruhe Shower Core Array Detector (KASCADE) experiment, which uses electromagnetic, muonic, and hadronic particle detectors focused on studying air showers in the energy range around the knee (10^{15} to 10^{17} eV). These data provide evidence for a transition from light nuclei to heavier ones, with a hint of iron becoming dominant just above 10^{17} eV (*4*). A high-energy extension named KASCADE-Grande will reach 10^{18} eV to test this indication in the very near future. Further in the future, low-energy extensions of UHECR observatories such as the Auger Muons and Infill for the Ground Array (AMIGA) and High-Elevation Auger Tele-

Department of Astronomy and Astrophysics, Enrico Fermi Institute, Kavli Institute for Cosmological Physics, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637, USA. E-mail: olinto@kicp.uchicago.edu

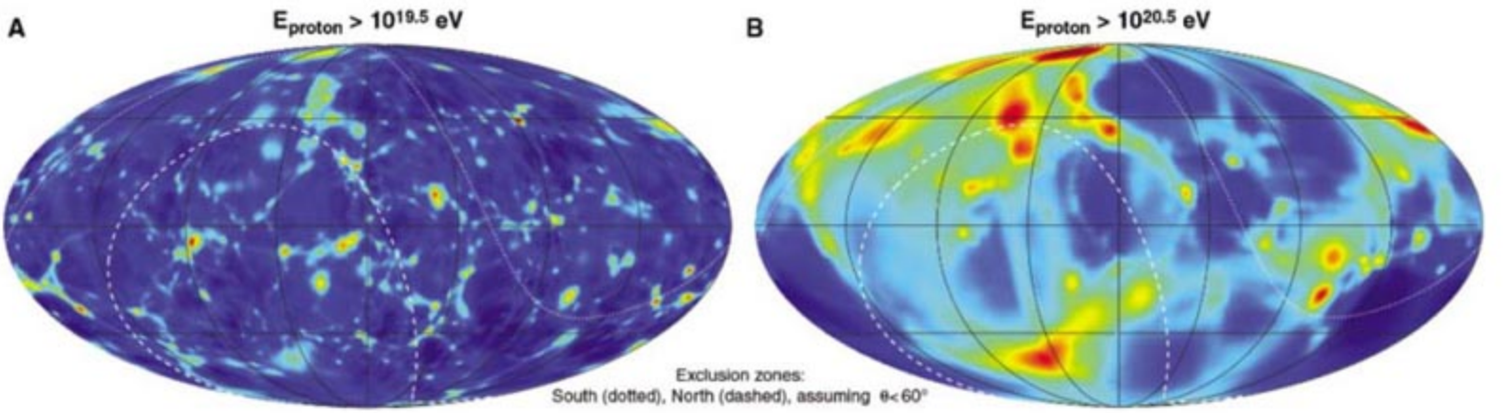


Fig. 1. Sky maps of predicted arrival directions of UHECR with energies of about (**A**) $10^{19.5}$ eV and (**B**) $10^{20.5}$ eV, assuming sources correlate with the dark-matter distribution. The map is a density contrast of arrival events in a log scale ranging from $10^{-2.2}$ (in blue) to 1 in red. [Image courtesy of (*8*)]

scope (HEAT) projects at the Pierre Auger Observatory and the Telescope Array Low-Energy (TALE) extension project of the Telescope Array plan to bridge the study of cosmic rays from just above the knee to the ankle region and beyond.

Among the many proposals for the origin of cosmic rays, the leading candidate for the acceleration of galactic cosmic rays is stochastic shock acceleration in supernova remnants (SNRs), based on a first-order Fermi acceleration mechanism that evolved from a 1949 proposal of Enrico Fermi. SNR shock acceleration naturally generates a power law spectrum of about the right slope, has the necessary energetic requirements, and may explain the observed composition trends (*5*). A clear confirmation of this picture is still lacking, but indirect support for this model has recently accumulated. Chandra satellite x-ray images of SNRs Tycho and SN1006 have indicated that relativistic electrons gain energy in a very thin region at the boundary of SNRs, where magnetic fields reach several hundred microgauss (*6*). More recently, the High-Energy Stereoscopic System (HESS) Imaging Atmospheric Cherenkov Telescope (IACT) array has produced the first images of SNRs in TeV gamma rays. Most notable is the image of RX J1713.7-3946 (*7*), in which it is clear that SNR shells emit TeV gamma rays, consistent with the evidence that they are produced by the decay of neutral pions at the sites of high-energy hadronic interactions. To clearly discriminate between the smoking gun of hadronic acceleration and the production of TeV gamma rays by electronic inverse Compton scattering, it is important to extend the spectrum of RX J1713.7-3946 below the energy threshold of HESS. This should be achieved in the near future by the Gamma-Ray Large-Area Space Telescope (GLAST) satellite, which is scheduled to launch in 2007.

Even if shock acceleration in supernova remnants is responsible for accelerating cosmic rays up to the knee, it is hard to imagine that this mechanism can reach much beyond ~10^{16} eV. At the highest energies, even more powerful sources seem to be required. In addition, as the energy of the primary cosmic ray increases, the effect of the galactic magnetic field on the particle trajectory decreases. As cosmic rays reach energies of ~10^{19} eV and above, trajectories should point back to cosmic ray sources—i.e., cosmic ray astronomy should become possible. Thus far, observations show an isotropic distribution of arrival directions up to the highest energies observed. With no indication of the galactic plane or other nearby structures, this isotropy argues for an extragalactic origin for the highest-energy particles. If UHECRs (above 10^{18} eV) originate in extragalactic sources distributed equally throughout the universe, the distribution of arrival directions in the sky will be isotropic to first order, given that protons of 10^{18} eV can traverse the entire universe unimpeded. As observations of cosmic rays from 10^{19} to 10^{20} eV begin to accumulate in statistics, the effect of the GZK feature should induce a marked change in the distribution of arrival directions of UHECRs. Instead of an isotropic universe, we should see the anisotropic galaxy distribution in our local 10- to 100-Mpc volume.

Fig. 2. A water Cherenkov tank of the Auger Observatory in the Argentinean Pampa Amarilla.

Figure 1 shows the predicted change in anisotropies in the arrival-direction distribution of UHECRs as the observed energy changes from $10^{19.5}$ to $10^{20.5}$ eV. These figures were produced (*8*) assuming that UHECR sources trace the dark-matter distribution in the universe. The contrast at the highest energies is only a factor of 2, which underscores the challenge of charged-particle astronomy: the observation of small anisotropies as the cosmic ray flux reaches below 1 particle per km^2 per century. Newly constructed and future UHECR observatories will answer this challenge by covering areas of 3000 km^2, such as the southern site of the Pierre Auger Observatory, and even larger areas, as proposed for the Northern site and space missions.

UHECRs are detected by means of two main techniques: ground arrays (of scintillators or water Cherenkov tanks) and fluorescence telescopes. Ground arrays sample the extensive air shower as the secondaries reach the ground. The largest arrays to explore UHECRs include Haverah Park (1967 to 1987), Sydney University Giant Air-Shower Recorder (SUGAR) (1968 to 1979), Yakutsk (1991 to present), and the largest before the Pierre Auger Observatory, the Akeno Giant Air-Shower Array (AGASA). The 111 surface detectors of AGASA covered 100 km^2 and operated for just over a decade (1990 to 2004), reaching an exposure of 1.6×10^3 L during the project's lifetime (the unit of exposure, L = 1 km^2 sr year, is named after J. Linsley). An alternative technique based on atmospheric fluorescence was pioneered by the Fly's Eye detector, which in 1991 observed an event with energy 3×10^{20} eV. The fluorescence technique was further developed by the High-Resolution Fly's Eye

(HiRes) experiment, which reached an exposure slightly higher than the AGASA exposure in the recent past. These observatories detect the fluorescence of nitrogen molecules in the atmosphere as the shower develops above the ground. Mirrors focus the fluorescent ultraviolet light onto photomultiplier tubes that record the fast-moving shower pattern in the atmosphere. This technique, unlike ground arrays, can observe the shower maximum directly. However, it has a low duty cycle that works best during clear moonless nights.

Since the prediction of the GZK feature in 1966, progress in the field has been hindered by the experimental challenge of reaching exposures greater than 10^4 L. The HiRes and AGASA observatories gave conflicting results on the existence of the GZK feature (*9*–*11*), hampered by the low statistics and systematic discrepancies in the energy scale (*12*). The exposure challenge will be faced soon by the completion of the southern site of the Pierre Auger Observatory (*13*). When completed in 2007, the Southern Auger Observatory in the Mendoza province of Argentina will cover 3000 km^2 in a ground array of water Cherenkov detectors (Fig. 2) overlooked by four fluorescence telescope sites. This first hybrid detector uses the strengths of both techniques: the high statistics and geometrical aperture of the ground array with the high-quality reconstruction of 10% of showers observed with the fluorescence telescopes. Auger South has been accumulating data during construction and should reach 10^4 L by 2008. In this exposure range, Auger South will provide high statistics measurement of the spectral features together with composition estimates between $10^{17.5}$ and 10^{20} eV. In addition to resolving the conflict over the shape of the UHECR spectrum around the GZK feature, Auger South will also help determine the transition from galactic to extragalactic cosmic rays expected to occur between 10^{17} and 10^{19} eV. The precise spectral and composition measurement over this wide range of energies will constrain the injection spectrum and composition of proposed UHECR sources as well as the effect of source distribution and magnetic fields on the propagation of UHECRs from source to Earth.

Auger South will explore the 10^4-L exposure range during most of its lifetime and should make a precise measurement of the long-awaited "cosmologically meaningful termination." In addition, neutrino telescopes such as the Antarctic Impulsive Transient Antenna (ANITA), Super Radio Ice Cherenkov Experiment (SuperRICE), IceCube, Low-Frequency Array (LOFAR), and possibly the Square-Kilometer Array (SKA), will explore the predicted neutrino flux from the interactions of UHECRs with the CMB that give rise to the GZK feature. The multimessenger approach to the origin of UHECRs will establish their origin as extragalactic and begin to focus on possible sources. A key ingredient in the unveiling of UHECR sources will be the detection of anisotropies in the arrival distribution of UHECRs (Figs. 1 and 2); a new generation of observatories is now being planned to achieve this goal. On the ground, the proposed Northern site of the Auger Observatory would cover an area of 4000 square miles (10,370 km^2) to reach 10^5 L in the next decade. In space, fluorescence telescopes are being planned to look down on Earth from the International Space Station or from free-flying dedicated satellites. The era of 10^5 L will open the new field of charged-particle astronomy.

References and Notes

1. K. Greisen, *Phys. Rev. Lett.* **16**, 748 (1966).
2. G. T. Zatsepin, V. A. Kuzmin, *Sov. Phys. JETP Lett.* **4**, 781966 (1966).
3. T. K. Gaisser, T. Stanev, *Nucl. Phys. A.* **777**, 98 (2006).
4. T. Antoni *et al.*, *Astropart. Phys.* **24**, 1 (2005).
5. A. M. Hillas, in *Cosmology, Galaxy Formation and Astroparticle Physics on the Pathway to the SKA*, H.-R. Klöckner, M. Jarvis, S. Rawlings, Eds., Oxford, UK, 10 to 12 April 2006 (Univ. of Oxford Press, Oxford, UK, in press); available online (http://arxiv.org/pdf/astro-ph/0607109).
6. H. J. Völk, E. G. Berezhko, L. T. Ksenofontov, *Astron. Astrophys.* **433**, 229 (2005).
7. F. Aharonian *et al.*, *Astron. Astrophys.* **449**, 223 (2006).
8. E. Armangaud, personal communication.
9. N. Hayashida *et al.*, *Phys. Rev. Lett.* **73**, 3491 (1994).
10. T. Abu-Zayyad *et al.* (High Resolution Collaboration), *Astropart. Phys.* **23**, 157 (2005).
11. R. U. Abbasi *et al.* (High Resolution Collaboration), *Phys. Rev. Lett.* **92**, 151101 (2004).
12. D. DeMarco, P. Blasi, A. V. Olinto, *Astropart. Phys.* **20**, 53 (2003).
13. J. Abraham *et al.* (Pierre Auger Collaboration), *Nucl. Instr. Methods* **A523**, 50 (2004).

14. I acknowledge funding by the NSF through grants NSF PHY-0457069 and by the Kavli Institute for Cosmological Physics under NSF PHY-0114422 at the University of Chicago.

10.1126/science.1136640

PERSPECTIVE

The Very-High-Energy Gamma-Ray Sky

Felix Aharonian[1,2]

Over the past few years, very-high-energy gamma-ray astronomy has emerged as a truly observational discipline, with many detected sources representing different galactic and extragalactic source populations—supernova remnants, pulsar wind nebulae, giant molecular clouds, star formation regions, compact binary systems, and active galactic nuclei. It is expected that observations with the next generation of stereoscopic arrays of imaging atmospheric Cherenkov telescopes over a very broad energy range from 10^{10} to 10^{15} electron volts will dramatically increase the number of very-high-energy gamma-ray sources, thus having a huge impact on the development of astrophysics, cosmology, and particle astrophysics.

It has been said that very-high-energy (VHE) gamma-rays—photons with energy in excess of 100 billion eV (*1*)—represent the "last window" onto cosmic electromagnetic radiation. They are copiously produced, thanks to various electromagnetic and hadronic interactions, in nature's machines, cosmic TeVatrons and PeVatrons, which are capable of accelerating electrons, protons, and nuclei to TeV and PeV energies. Unlike charged particles, gamma-rays freely propagate through the intergalactic radiation and magnetic fields across most of the universe. Finally, they are detectable by space-borne or ground-based detectors. These three features make very-high-energy gamma-rays unique carriers of astrophysical and cosmological information about the most energetic and violent processes in the universe.

Gamma-ray astronomy addresses a diverse range of topics in modern astrophysics and particle astrophysics, including (i) acceleration and radiation processes in extreme conditions, in particular in relativistic outflows like jets and winds formed in the vicinity of black holes and pulsars; (ii) the origin of galactic and extragalactic cosmic rays; (iii) the nature of nonthermal transient phenomena such as gamma-ray bursts; (iv) cosmology, by probing the cumulative extragalactic background light that contains information about the history of formation of galaxies and the first stars; and (v) fundamental physics, including the indirect search for dark matter and signals from primordial black holes.

Earth's atmosphere is not transparent to gamma-rays; therefore, an ideal detector would be located in space. However, space platforms offer limited detection areas, effectively constraining the study of weak cosmic gamma-ray fluxes to energies below 100 GeV. At higher energies, an alternative method of detection of cosmic gamma-rays becomes available, based on the registration of secondary showers produced by interactions of primary gamma-rays with Earth's atmosphere, seen either directly or through their Cherenkov radiation. Because the speed of ultrarelativistic electrons exceeds the speed of light in the atmosphere, these electrons produce an ~1° cone of blue Cherenkov light that forms a pool on the ground with a radius of about 120 m. The Cherenkov signal of air showers is very faint and brief; the flash lasts only a few nanoseconds. Consequently, Cherenkov telescopes must have large (>>1 m^2)

[1]Dublin Institute for Advanced Studies, Dublin, Ireland. [2]Max Planck Institut fur Kernphysik, Heidelberg, Germany. E-mail: Felix.Aharonian@mpi-hd.mpg.de

optical reflectors to image the Cherenkov light onto a very fast multipixel camera sensitive to the blue light with a typical pixel size of 0.1° to 0.2° and a field of view of several degrees. The total number of photons collected in the resulting image is a measure of energy, the orientation of the image correlates with the arrival direction of the gamma-ray, and the shape of the image contains information about the origin of the primary particle (a proton or gamma-ray). These three features, coupled with the huge (as large as 0.1 km^2) detection area, comprise the basis of the Imaging Atmospheric Cherenkov Telescope (IACT) technique.

The first reliable VHE gamma-ray signal from an astronomical object, the Crab Nebula, was detected using the IACT technique in the late 1980s by the Whipple 10-m-diameter telescope located on Mt. Hopkins, Arizona (*2*). Over the next 15 years, major efforts to detect gamma-rays were made by the Cherenkov Array at Themis (CAT), Collaboration of Australia and Nippon (Japan) for a Gamma-Ray Observatory in the Outback (CANGAROO), High-Energy Gamma-Ray Astronomy (HEGRA), Whipple, and some other groups. However, they only managed to detect 10 or so VHE gamma-ray sources, some tentatively. So, despite several notable results, in particular the discovery of gamma-rays from blazars (*3*), these efforts did not present a huge breakthough. More sensitive detectors were needed badly.

In the mid 1990s, the concept of stereoscopic arrays, consisting of two or more 10-m-diameter class telescopes observing the flashes simultaneously from different directions, was recognized as the most promising approach that can facilitate dramatic improvement in the sensitivity and push the detection threshold down to 100 GeV (*4*). Although the power of the stereoscopic approach was convincingly demonstrated by the HEGRA system of small aperture telescopes, it was the High-Energy Stereoscopic System (HESS) that elevated the status of the field to a level of a truly observational (astronomical) discipline. HESS, an array of four 13-m-diameter IACTs equipped with an ~5° field of view imagers, was completed in 2004 (Fig. 1). It covers a broad energy band from 100 GeV to 100 TeV with an angular resolution of a few arc minutes and minimum detectable energy flux approaching 10^{-13} erg cm^{-2} s^{-1}. Whereas HESS observes sources mainly from the Southern Hemisphere of the sky, Major Atmospheric Gamma-Ray Imaging (MAGIC)—a single very large Cherenkov telescope—targets the Northern Hemisphere (Fig. 1). Soon the Very Energetic Radiation Imaging Telescope Array System (VERITAS), a new stereoscopic array consisting of four IACTs, will start taking data from Southern Arizona.

Presently, several galactic and extragalactic source classes are established as TeV gamma-ray emitters (see Fig. 1). One of the remarkable achievements of HESS was the discovery of shell-type structures of young supernova remnants (see Fig. 2), in particular of the object RXJ1713.7-3946 (*5*, *6*), which was earlier reported as a TeV gamma-ray source by the CANGAROO collaboration (*7*). This result supports the early theoretical predictions that galactic cosmic rays must have deep links to supernova remnants, namely, that they are accelerated by shocks in the shells of material lost during the supernova explosion.

HESS has also revealed that many young pulsars are surrounded by extended regions of VHE gamma-ray emission. Some show an energy-dependent morphology (*8*), such that the source size reduces as the photon energy increases (see Fig. 2). This can be explained by the energy losses of electrons and strongly supports the paradigm that electrons are accelerated to 100 TeV energies and beyond at the site of the termination of the cold ultrarelativistic pulsar wind.

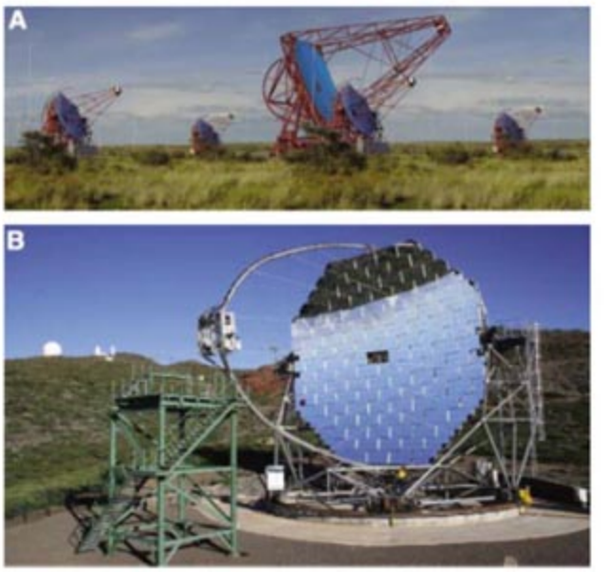


Fig. 1. HESS and MAGIC. **(A)** HESS is a stereoscopic array of four 13-m-diameter Cherenkov telescopes located in Namibia. The central telescope is an artistic view of the new 28-m-diameter dish, presently under construction. [Image, W. Hofmann] **(B)** MAGIC is a single 17-m Cherenkov telescope located on the Canary Island of La Palma. In 2007, it will be accompanied by a second similar telescope. This will allow observations of gamma-rays sources in a stereoscopic mode. [Image, P. Sawallisch]

If a particle accelerator is located in a binary system with a luminous optical star, the interactions of accelerated electrons with the optical starlight or with the dense stellar wind proceed on time scales of hours or even less. Thus, such binary systems allow continuous watch of the complex acceleration and magnetohydrodynamic processes such as the creation and termination of relativistic outflows related to the compact object. This may be a "cold" pulsar wind in the case of a neutron star or a "hot" jet in the case of a black hole. So far, three compact binary systems have been detected by the HESS (*9*–*11*) and MAGIC telescopes (*12*). The so-called Microquasar LS 5039, a binary star system where one component is a black hole, shows a strictly periodic component, which implies that the source behaves as a "TeV clock" with a period of 3.908 ± 0.002 days, which perfectly coincides with the known orbital period of the system (*11*).

Although gamma-rays from discrete objects reveal the locations of cosmic accelerators, one should expect also a diffuse component of radiation caused by interactions of relativistic particles, which escape their production sites, with the surrounding dense gas regions like giant molecular clouds (GMCs). The HESS observations indeed revealed TeV gamma-ray emission which correlates with several distinct GMCs in the central 200 parsec region of our Galaxy (*13*). The gamma-ray map of this region indicates an inhomogeneous spatial distribution of the runaway protons, which can be explained by the high activity of the particle acceleration in the past, related, for example, to the compact radio source Sgr A* or to a recent supernova explosion in the galactic center. Sgra A*—presumably a supermassive black hole (SMBH) located in the dynamical center of our Galaxy—can be responsible also for the compact TeV gamma-ray source detected by the Whipple, CANGAROO, HESS and MAGIC groups, although some other explanations, including the hypothetical Dark Matter Halo of the Galaxy, cannot be excluded.

More compelling evidence for production of gamma-rays in SMBHs recently was obtained from Giant Radiogalaxy M87. The detected variability of TeV gamma-ray emission on time scales of days implies that we "see" gamma-rays arriving from regions located in the vicinity of a 3×10^9 solar mass black hole (*14*).

SMBHs, the powerhouses of active galactic nuclei (AGN), play a key role in production of VHE gamma-rays observed from AGN. The gamma-ray horizon of the universe, the most distant observable region, is determined by gamma-ray interactions with the diffuse extragalactic background light (EBL); at very high energies, it is limited to distances of only several hundred megaparsecs. That is why the first extragalactic TeV gamma-ray sources were dominated by relatively nearby blazars—AGN with jets directed toward Earth. Although the effect of relativistic Doppler boosting provides orders of magnitude enhancement of the gamma-ray flux, the nearby location of these sources is crucial to minimize absorption of gamma-rays and thus make feasible their detection. With reduction of the energy threshold of detectors down to <100 GeV, one should expect a substantial increase in

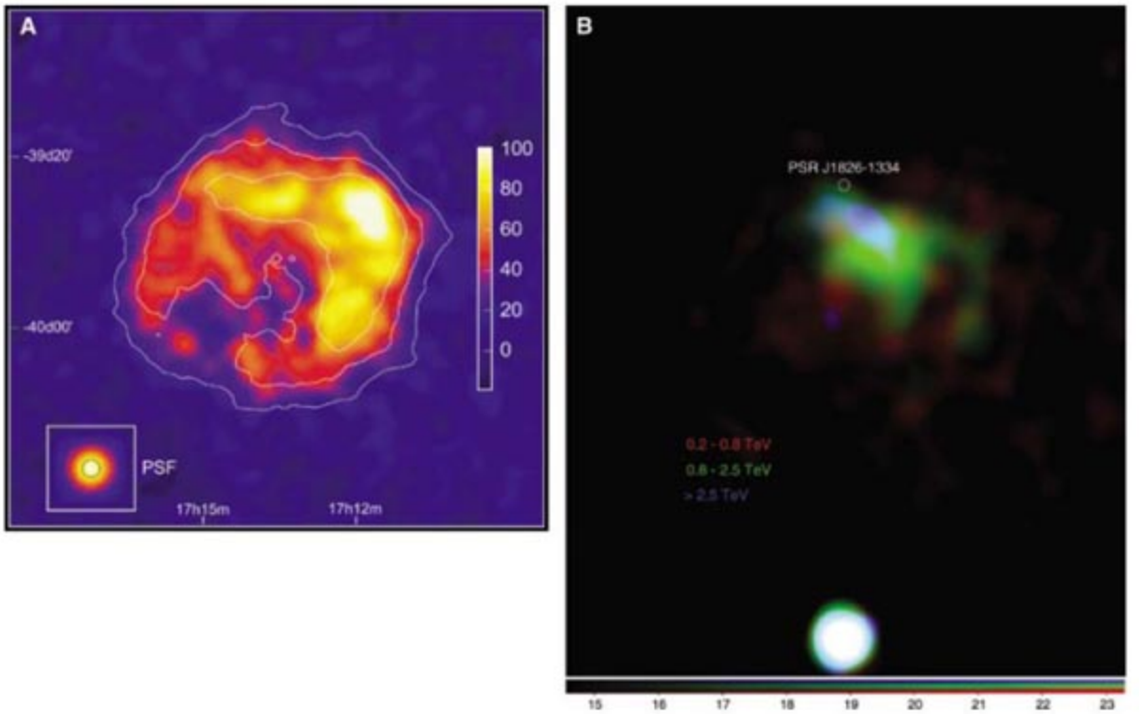


Fig. 2. (**A**) The gamma-ray image of the young SNR RX 1713.7-3946 obtained with the HESS telescope array. The shell-type morphology is clearly seen (*5*). [Image, D. Berge] (**B**) The gamma-ray images of the extended gamma-ray source HESS J1825-137 shown for different energy bands: below 1 TeV (red), between 1 and 2.5 TeV (green), above 2.5 TeV (blue). The source most likely associates with the pulsar PSR J1826-1334, the location of which is indicated by the white point. Gamma-ray production proceeds mainly through the inverse Compton scattering of these electrons on photons of the 2.7 K cosmic microwave background radiation. Because the latter fills every corner of the cosmos, the spatial and energy distributions of electrons can be derived from VHE gamma-ray data unambiguously and with very high precision—a unique case in astrophysics when the nonthermal particle distributions are obtained without any additional assumptions. The bright point-source to the south is the microquasar LS5039. [Image, S. Funk]

the numbers of extragalactic objects detected. Intensive searches conducted by HESS and MAGIC collaborations over the past 2 years have doubled the number of known TeV blazars. Some are quite distant, reaching redshifts of $z = 0.20$. This result was used to derive a robust upper limit on the EBL flux at optical/near-infrared wavelengths so as to constrain cosmological models concerning the formation and evolution of galaxies and the first stars (*15*).

Planning of the next generation of IACT arrays (*16*) has two objectives: (i) an order-of-magnitude improvement of the flux sensitivity in the standard (0.1 to 10 TeV) energy regime and (ii) an aggressive expansion of the energy domain of IACT arrays in both directions, down to 10 GeV and up to 1 PeV.

If one limits the energy to ~100 GeV, the performance of the telescope arrays can be predicted with confidence. Namely, a sensitivity well below 10^{-13} erg cm^{-2} s^{-1} and angular resolution of 1 to 2 arc min can be achieved by a stereoscopic array consisting of a very large number (up to 100) of 10-m-diameter class IACTs. One may predict, based on the extrapolation of the HESS results, that such an instrument will discover and resolve hundreds, or perhaps even thousands, of galactic TeV sources. On the other hand, such an array would gain a lot if the energy threshold can be reduced to 30 GeV. This would considerably increase the number of scientific objectives, in particular increase the distance range of detectable extragalactic objects up to redshifts of $z = 1$, as well as considerably improve the flux sensitivity around 100 GeV. This can be achieved by somewhat larger 15-m-diameter class telescopes, installed at quite high altitudes of 3 to 4 km above sea level. The construction of such a powerful detector could be completed on relatively short time scales because it would be based on current technologies.

Further reduction of the energy threshold down to 10 GeV or even less is possible but requires a different approach: operation of 30-m-diameter class telescopes in a robotic regime at extremely high altitudes of 5 km above sea level—for example, on the Atacama Large Millimeter Array (ALMA) site (*17*) and design of high quantum efficiency focal plane imagers. The energy range from several GeV to 30 GeV has very specific astrophysical and cosmological objectives: exploration of highly variable nonthermal phenomena, in particular in the remote universe at redshifts of $z = 5$, as well as in compact galactic objects like microquasars. The successful realization of such a gamma-ray timing explorer, hopefully during the lifetime of the Gamma-Ray Large-Area Space Telescope (GLAST) mission (*18*), would be a great achievement for gamma-ray astronomy.

Finally, it is important to develop a ground-based technique allowing simultaneous coverage of a substantial fraction (1 steradian or more) of the sky. The most realistic approach uses very large water Cherenkov detectors installed at altitudes of ~4 km (*19*). The feasibility of this technique has been convincingly demonstrated by the Milagro collaboration. The prospect of exciting discoveries of yet unknown VHE transient phenomena in the universe fully justifies the efforts toward the construction of a large field-of-view ground-based gamma-ray detector(s). These instruments will be complementary to GLAST and the future large-volume (km^3-scale) high-energy neutrino detectors.

References and Notes

1. The standard energy units often used in particle astrophysics are MeV (10^6 eV), GeV (10^9 eV), TeV (10^{12} eV), and PeV (10^{15} eV).
2. T. Weekes *et al.*, Whipple Collaboration. *Astrophys. J.* **342**, 379 (1989).
3. M. Punch *et al.*, Whipple Collaboration. *Nature* **358**, 477 (1992).
4. F. A. Aharonian, C. W. Akerlof, *Annu. Rev. Nucl. Part. Sci.* **47**, 273 (1997).
5. F. A. Aharonian *et al.*, HESS collaboration. *Nature* **432**, 75 (2004).
6. F. A. Aharonian *et al.*, *Astron. Astrophys.* **449**, 223 (2006).
7. R. Enomoto *et al.*, CANGAROO collaboration. *Nature* **416**, 823 (2002).
8. F. A. Aharonian *et al.*, HESS collaboration, *Astron. Astrophys.* **460**, 743 (2006).
9. F. A. Aharonian *et al.*, HESS collaboration. *Astron. Astrophys.* **442**, 1 (2005).
10. F. A. Aharonian *et al.*, HESS collaboration. *Science* **309**, 746 (2005).
11. F. A. Aharonian *et al.*, HESS collaboration *Astron. Astrophys.* **460**, 365 (2006).
12. J. Albert *et al.*, MAGIC collaboration. *Science* **312**, 1771 (2006).
13. F. A. Aharonian *et al.*, HESS collaboration. *Nature* **439**, 695 (2006).
14. F. A. Aharonian *et al.*, HESS collaboration. *Science* **314**, 1424 (2006).
15. F. A. Aharonian *et al.*, HESS collaboration. *Nature* **440**, 1018 (2006).
16. B. Degrange, G. Fontaine, Eds., *Proceedings of the Conference Toward a Network of Atmospheric Cherenkov Detectors VII* (Ecole Polytechnique, Palaiseau, France, 2005).
17. ALMA is a project of an array of 12-m-diameter antennas for studies in the submillimeter range to be installed on a plateau at 5-km elevation at Llano de Chajnantor, Chile.
18. GLAST is NASA's major spaceborne gamma-ray mission for study of the MeV/GeV sky. It will be launched in 2007.
19. G. Sinnis, High Altitude Water Cherenkov telescope (HAWC) Collaboration, *AIP Conference Proceedings*, vol. 745, p. 234 (American Institute of Physics, New York, 2005).

10.1126/science.1136395

Cross-Species Identification of Mendel's *I* Locus

Ian Armstead,[1]* Iain Donnison,[1] Sylvain Aubry,[2] John Harper,[1] Stefan Hörtensteiner,[2] Caron James,[1] Jan Mani,[2] Matt Moffet,[3] Helen Ougham,[1] Luned Roberts,[1] Ann Thomas,[1] Norman Weeden,[3] Howard Thomas,[1] Ian King[1]

We identified gene homologs in *Pisum sativum* (pea), *Arabidopsis thaliana*, and *Festuca pratensis* (meadow fescue), mutations of which partially disable plant senescence. The biochemical properties and map location of this gene in pea indicate that it is the same locus that determines yellow (*I*) and green (*i*) cotyledon color, as originally described by Mendel in his seminal paper (*1*). Staygreen (the gene *sgr*), the indefinite retention of greenness in senescing leaves and cotyledons, is a genetic variant that interferes with the normal expression of senescence. Biochemical (*2*, *3*) (Fig. 1A) and genetic analyses of *sgr* from pea and *F. pratensis* (*4*) show that it segregates as a discrete phenotype with equivalents in maize, sorghum, and rice (*5*, *6*). By using the *Lolium-Festuca* gene introgression system and extrapolating from genetic synteny, we show that *F. pratensis sgr* on chromosome (C)5 maps to a position corresponding to *sgr* on rice C9 (*4*, *6*, *7*). Fine mapping in *Festuca* delimits this to a region equivalent to circa 200 kb of the rice genome containing 30 annotated gene models, including a predicted senescence-inducible chloroplast-specific stay-green protein (*4*, *7*, *8*).

We investigated the temporal and tissue-specific expression patterns of the most similar *Arabidopsis* gene model, At4g22920, through the Genevestigator Meta-Analyzer database (*9*). These patterns indicate that At4g22920 is up-regulated in days 45 to 50 during maximal senescence in the *Arabidopsis* life cycle (fig. S1A) and that the strongest expression of this gene occurs in senescent leaves, although it is also detected in seeds, petals, and sepals. RNA interference (RNAi) was also used to silence the expression of At4g22920 in *Arabidopsis*, resulting in plants with leaves exhibiting prolonged chlorophyll retention upon dark incubation (Fig. 1B), phenotypically equivalent to *sgr F. pratensis*. Thus, alleles of *F. pratensis* and rice homologs of *Arabidopsis* At4g22920 likely underlie the monocot stay-green phenotypes.

In order to explore the parallels between Mendel's pea cotyledons and the stay-green leaf phenotypes in *Arabidopsis*, *Festuca*, and rice, we genetically mapped the pea homolog of *sgr* in two different pea populations segregating for cotyledon color polymorphism (*I* and *i*). No recombination between *I* and *sgr* was observed in 104 progeny, reflecting a theoretical maximum distance between *sgr* and *I* of 0.48 cM (fig. S1B). Northern analysis in pea indicates that *sgr* has reduced expression in mutant senescing leaves when compared with that of wild type (Fig. 1C).

This report characterizes a rare example of an unbroken chain from gene sequence to trait through biochemical, cell biological, and physiological definitions of phenotype, in which genomics and functional analysis in model species have played a critical role. This study results in the identification of a gene that plays a fundamental role in chlorophyll catabolism during plant senescence. In addition, it suggests that the cotyledon color trait described by Mendel reflects allelic variation in a pea gene, homologs of which are responsible for the stay-green phenotype in both dicots and monocots.

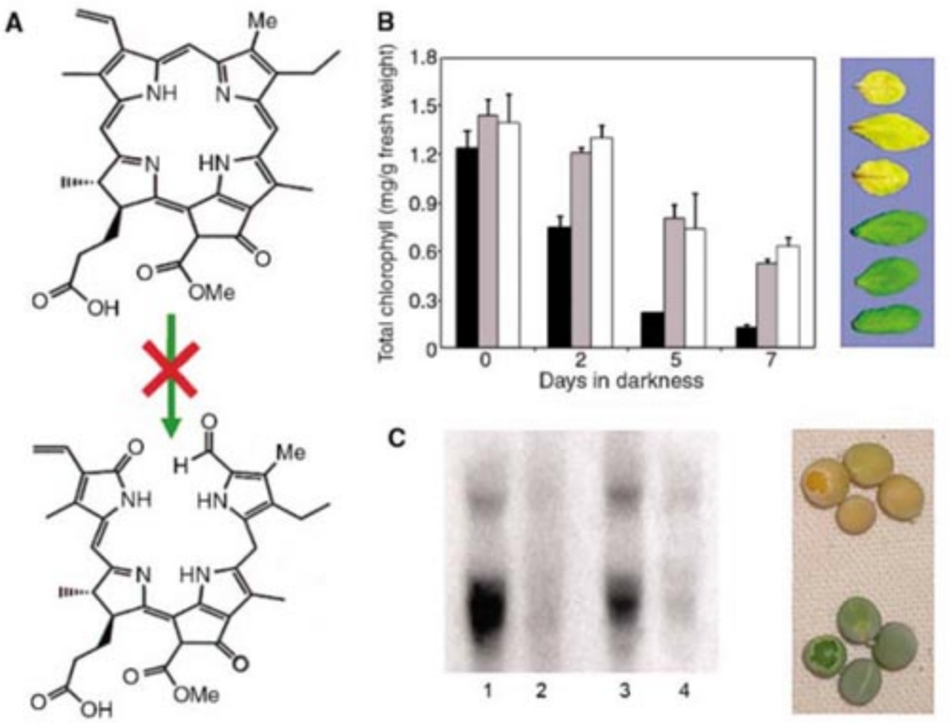


Fig. 1. (**A**) In mutant and silenced genotypes, inhibition (X) of the ring-opening step between (top) pheophorbide and (bottom) red chlorophyll catabolite in the chlorophyll breakdown pathway (*3*) leads to retention of greenness in senescing leaves. (**B**) RNAi silencing of *Arabidopsis* At4g22920 causes a stay-green phenotype. (Left) Degradation of chlorophyll in control (black) and two independently RNAi-silenced genotypes (gray and white) during dark-induced, detached leaf senescence. Error bars indicate standard deviation. (Right) Wild-type (top) and stay-green (bottom) leaf phenotypes of *Arabidopsis* after 5 days of dark-induced, detached-leaf senescence. (**C**) (Left) Northern analysis using RNA extracted from senescing leaves of wild-type (lanes 1 and 3) and stay-green (lanes 2 and 4) pea plants. (Right) Wild-type (top) and stay-green (bottom) pea cotyledons illustrating Mendel's *I* and *i* phenotypes, respectively.

References and Notes

1. G. Mendel, *Verh. Naturforsch. Ver. Brünn.* **4**, 3 (1866).
2. H. Thomas, M. Schellenberg, F. Vicentini, P. Matile, *Bot. Acta* **109**, 3 (1996).
3. S. Hörtensteiner, *Annu. Rev. Plant Biol.* **57**, 55 (2006).
4. I. P. Armstead *et al.*, *New Phytol.*, **172**, 592 (2006).
5. H. Thomas, C. M. Smart, *Ann. Appl. Biol.* **123**, 193 (1993).
6. K. W. Cha *et al.*, *Theor. Appl. Genet.* **104**, 526 (2002).
7. Materials and methods are available as supporting material on *Science* Online.
8. Institute for Genomic Research (TIGR), gene model LOC_Os09g36200, www.tigr.org.
9. Eidgenössische Technische Hochschule Zurich, www.genevestigator.ethz.ch/.
10. This work was supported by the Biotechnology and Biological Sciences Research Council (UK), the Swiss National Science Foundation (3100A0-105389), and Cooperative State Research, Education, and Extension Service, U.S. Department of Agriculture, for Hatch Project MONB00235.

Supporting Online Material

www.sciencemag.org/cgi/content/full/315/5808/73/DC1
Materials and Methods
Fig. S1
Table S1
References

24 July 2006; accepted 16 October 2006
10.1126/science.1132912

[1]Institute of Grassland and Environmental Research, Aberystwyth SY23 3EB, UK. [2]Institute of Plant Sciences, University of Bern, Altenbergrain 21, CH-3013 Bern, Switzerland. [3]Department of Plant Sciences and Plant Pathology, Montana State University, Bozeman, MT 59717, USA.

*To whom correspondence should be addressed. E-mail: ian.armstead@bbsrc.ac.uk

Atom Interferometer Measurement of the Newtonian Constant of Gravity

J. B. Fixler,[1] G. T. Foster,[2] J. M. McGuirk,[3] M. A. Kasevich[1]*

We measured the Newtonian constant of gravity, G, using a gravity gradiometer based on atom interferometry. The gradiometer measures the differential acceleration of two samples of laser-cooled Cs atoms. The change in gravitational field along one dimension is measured when a well-characterized Pb mass is displaced. Here, we report a value of $G = 6.693 \times 10^{-11}$ cubic meters per kilogram second squared, with a standard error of the mean of $\pm 0.027 \times 10^{-11}$ and a systematic error of $\pm 0.021 \times 10^{-11}$ cubic meters per kilogram second squared. The possibility that unknown systematic errors still exist in traditional measurements makes it important to measure G with independent methods.

The weak coupling of gravity compared with other forces makes precision gravity experiments difficult. This is manifested in the relatively poor knowledge of the Newtonian constant of gravity, G, compared with our understanding of other fundamental constants (*1*). The traditional torsion pendulum method for measuring G involves a well-characterized moving source mass that produces a torque on a test mass attached to a long fiber. Measurement of the test mass displacement, coupled with knowledge of the mechanics of the pendulum and of the source–test mass gravitational force, determines G. Other recent methods make use of a Fabry-Perot optical cavity (*2*), a flexure-strip balance (*3*), or a falling corner-cube gravimeter (*4*). The first direct precision measurement of G (*5*) determined the value of G to 1.1 parts per thousand (ppt), which remained the standard definition until 1942 (*6*), when the precision was increased to 0.45 ppt. During the past two decades, a number of high-precision measurements have been performed, but their discrepancies were larger than their standard deviations. Therefore, the accepted precision remained relatively unchanged (*7*, *8*). Recently, a few experiments have claimed to reach to <100 parts per million (ppm) (*9*–*11*).

The inherent difficulty of measuring G was evident in the change of the Committee on Data for Science and Technology (CODATA) definition from 1986 to 1998 (*1*, *8*), increasing the uncertainty to 1.5 ppt. Part of this increase came from an 83-ppm measurement in 1996 (*11*)—determined with the use of a dynamic fiberless torsion balance—that differed by 42 standard deviations from the CODATA value of G at the time. Questions were raised about the accuracy of other experiments as well. Also taken into account in the CODATA decision was a discovery of fiber twist anelasticity (*12*) in torsion balance experiments. Not until recently were the systematics of the experiment by Michaelis *et al.* (*11*) understood (*13*), resulting in the treatment of those results as outlier points.

We used quantum interference of atomic Cs to directly probe the gravitational scalar potential. The performance of instruments based on this technique meets or exceeds that of other state-of-the-art gravity (*14*) or gravity-gradient (*15*) devices. In this work, we used a gravity gradiometer to make a proof-of-principle measurement of G. We measured the differential acceleration of two laser-cooled ensembles of atomic Cs induced by a 540-kg Pb source mass precisely positioned between two vertically separated de Broglie wave gravimeters. With accurate knowledge of the atomic trajectories and the Pb source geometry and composition, we calculated the gravitationally induced phase shift in our atom interferometer and extracted a value for G. The accuracy was characterized with a thorough study of systematics that might influence our measurement. This method is loosely analogous to that of Schwarz *et al.* (*4*), who used macroscopic masses rather than interfering atomic wavepackets.

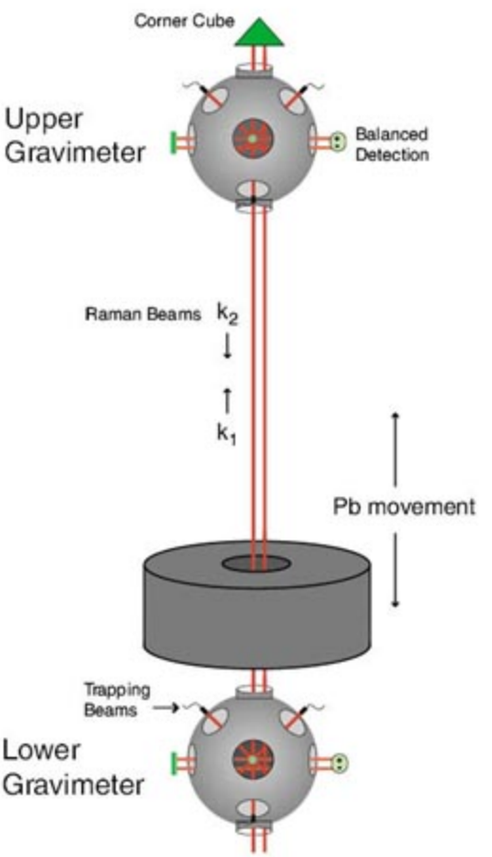


Fig. 1. Schematic of the experiment. The Raman beams propagate along a common vertical axis that contains both atomic ensembles (*20*).

Our gravity gradiometer consists of two gravimeters that operate by the light-pulse atom interferometry technique (Fig. 1) (*16*). The momentum recoil from the emission or absorption of photons by a Cs atom is used to coherently split and deflect the atomic wavepackets. A $\pi/2$ "splitter" pulse places an atom initially in the ground state with momentum p into a superposition of ground and excited states, $|g,p\rangle \rightarrow (|g,p\rangle + |e,p + \hbar k\rangle)/\sqrt{2}$, with the excited state gaining a photon recoil $\hbar k$ relative to the ground state part of the wavepacket ($k = 2\pi/\lambda$). A "mirror" π pulse drives an atom from the ground to the excited state, $|g,p\rangle \rightarrow |e,p + \hbar k\rangle$, imparting a photon recoil kick, or vice versa, which causes a stimulated emission of a photon and reduction of momentum. We applied a $\pi/2$-π-$\pi/2$ interferometer sequence with a pulse separation T (Fig. 2). The initial $\pi/2$ pulse separates the two wavepackets because of the difference in their momentum. The π pulse redirects the wavepacket momentum, causing the two components to overlap again at time $2T$, when the final $\pi/2$ pulse induces their interference. Momentum recoil creates different trajectories for the wavepackets that acquire a relative gravitationally induced atomic phase shift during the interferometer,

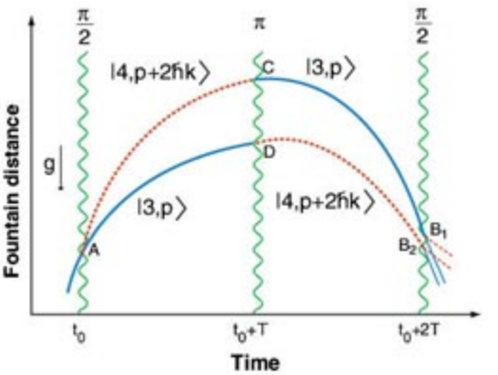


Fig. 2. Recoil space diagram of the atoms through the interferometer showing the separation (exaggerated) of the atomic wavepackets (*20*). A, initial $\pi/2$ pulse; B1, final $\pi/2$ pulse (upper trajectory); B2, final $\pi/2$ pulse (lower trajectory); C, π pulse (upper trajectory); D, π pulse (lower trajectory).

[1]Department of Physics, Stanford University, Stanford, CA 94305–4060, USA. [2]Department of Physics and Astronomy, City University of New York, Hunter College, New York, NY 10021, USA. [3]Department of Physics, Simon Fraser University, Burnaby, British Columbia, V5A 1S6, Canada.

*To whom correspondence should be addressed. E-mail: kasevich@stanford.edu

resulting in a sensitivity to accelerations. The total phase shift, $\Delta\phi_{Tot}$, is the sum of three components: the interaction of the atom with the light pulse, $\Delta\phi_{Laser}$; the quantum propagation phase accrued by each wavepacket over its trajectory, $\Delta\phi_{Path}$; and the wavepacket overlap, $\Delta\phi_{Separation}$ (*17*, *18*).

Operation of the gravity gradiometer has previously been described in detail (*15*, *19*, *20*). The two gravimeters are separated by 1.347 m, sharing a common vertical measurement axis. In each gravimeter chamber, we loaded Cs atoms into a magneto-optical trap (MOT) (*21*). The same trapping lasers used in the MOT then launched the atoms upward on a 12-cm ballistic trajectory. The interferometer pulses drove Doppler-sensitive two-photon optical (λ = 852 nm) Raman transitions (*22*) between the $F = 3$ and $F = 4$ hyperfine ground states. At the end of the interferometer, the probability that atoms will be in the $F = 4$ ground state for each gravimeter follows (*16*):

$$P_{|4\rangle} = \frac{1}{2}[1 - \cos(\phi_0 + \Delta\phi)] \quad (1)$$

where $\Delta\phi$ contains the gravitationally induced phase shift. Acousto-optic modulators in the Raman path are used to phase scan the fringe by changing ϕ_0. The transition probability is determined by detecting the atoms in both hyperfine states after the interferometer sequence with the use of a balanced, modulation-transfer technique (*23*).

Accelerations of the reference frame of the Raman laser field will result in phase shifts that are indistinguishable from gravitational accelerations by the Equivalence Principle. The gradiometer difference of simultaneous acceleration measurements in the same reference frame allows for rejection of vibrational noise. However, more than $(5 \times 10^{-7})g$ of environmental noise washes out the fringe contrast. We developed a technique to analyze the individual noisy data from the two gravimeters (*24*) that takes advantage of the idea that the two gravimeter signals parametrically describe an ellipse. Common phase noise in the two sinusoids distributes the data points around the ellipse but does not change the ellipticity. The ellipticity is proportional to the gravity-gradient phase difference. We used ellipse-specific fitting routines to extract the differential phase shift between the gravimeters without the need for the actively stabilized reference platform that would otherwise be required for individual fringe fitting.

The Pb source mass consisted of 20 stacked 2.5-cm-thick plates with an outer diameter of 35.3 cm and an inner bore with a diameter of 7.0 cm. The source was suspended between the gravimeters on a type 314 stainless steel platform, with a center bore, attached to an Al support frame. A stepper motor translated the frame vertically between the gravimeters, guided along a similar frame on Teflon pads. Two matched ball-screw jacks permitted high movement accuracy and repeatability, measured at 555 and 31 ppm, respectively.

We made a series of differential measurements of phase shifts induced by the Pb source at two positions, near the top of the lower gravimeter and 27.940 cm higher near the center of the gradiometer. The difference between the signal at the low and high positions provided a high level of rejection against possible systematic phase shifts that do not depend on the Pb position, including alternating-current Stark shifts, tidal field fluctuations, electronic offsets, and the Coriolis force. We periodically reversed the effective Raman propagation vector to gain further immunity to systematic phase offsets (*20*). A single data set typically lasted 7.6 hours.

During signal acquisition, the interferometer signal amplitudes, signal contrast, magnetic fields, laser-diode amplitudes, and room temperature were monitored. The gradient phase shift was observed to be independent of drifts in these parameters to within statistical deviations from a collection of data runs ($<5 \times 10^{-9}$ m/s^2). The phase was sensitive to the Raman lasers coming out of injection lock. This also produced a strong change in the signal amplitude or contrast, which allowed straightforward rejection during analysis.

The filtered data then was fit with the ellipse-specific routine. The output phases were separated into the two propagation phases and the two Pb positions. The phase data was then averaged for each position and propagation direction and combined to yield a mean chop phase difference for the data run. Standard deviations of the fitted subsets determined the phase uncertainty. A typical series of analyzed phase data is shown in Fig. 3. The modulation of the phase is clearly visible. The inset shows a typical 16-point fringe scan.

To determine a value for G from our differential chopped phase measurement, we modeled the expected signal from the Pb source potential. The model required accurate knowledge of the atomic trajectories in the two gravimeters, their respective distances to the Pb source distribution, and the Pb source characteristics (including Pb density and geometry). Given that the calculated total phase shift (*20*) exhibits a linear dependence on the initial trajectory position and velocity over the experimentally relevant range of initial conditions, the atom ensemble–averaged total phase shift is simply determined by the phase shift associated with the mean (measured) ensemble position and velocity. The ensemble position was determined with the use of resonance fluorescence and the velocity by time of flight (*20*).

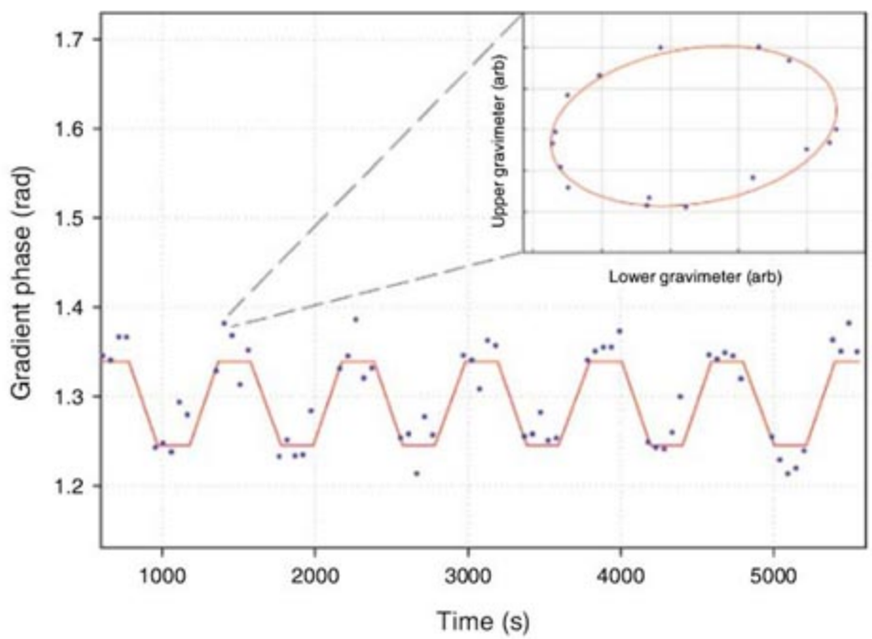


Fig. 3. A typical data sequence showing a modulation of the gradiometer phase output as the Pb source mass is displaced 27.940 cm from the top of the lower chamber. Data points are the fit phase of 16-point fringe scans. The single scan scatter was typically 35 mrad. The gravitational force from the Pb caused a differential acceleration of about $(-30 \times 10^{-9})g$ between the two interferometer signals. Solid trace shows the theoretical values. (Inset) A parametric plot of a typical 16-point scan with the ellipse-specific fit. arb, arbitrary units.

Table 1. Uncertainty limits.

Systematic	$\delta G/G$
Initial atom velocity	1.88×10^{-3}
Initial atom position	1.85×10^{-3}
Pb magnetic field gradients	1.00×10^{-3}
Rotations	0.98×10^{-3}
Source positioning	0.82×10^{-3}
Source mass density	0.36×10^{-3}
Source mass dimensions	0.34×10^{-3}
Gravimeter Separation	0.19×10^{-3}
Source mass density inhomogeneity	0.16×10^{-3}
Total	3.15×10^{-3}

Given the initial velocity and position, we determined the atom trajectories and then numerically solved for the total phase shift $\Delta\phi_{Tot} = \Delta\phi_{Laser} + \Delta\phi_{Path} + \Delta\phi_{Separation}$ using the exact potential of the source mass distribution (including the dominant contribution from the Pb as well as contributions from the stainless steel plate and support rods) and the second-order Taylor expansion of the Earth potential (*17*, *18*). Because the source mass potential depends linearly on G, the resulting total phase shift also depends on G. Thus, comparison of the calculated and observed shifts can be used to measure G. Evaluation of the term $\Delta\phi_{Path}$ involves an integral over the classical action for the calculated atomic trajectories. The term $\Delta\phi_{Separation}$ arises from the spatial separation of the two interfering wavepackets following the final $\pi/2$ pulse (this term is zero for uniform gravitational acceleration) and is obtained directly from the atomic trajectories. Finally, the term $\Delta\phi_{Laser}$ results from the light-pulse interactions and is determined by the phases of the laser fields evaluated at the semiclassical (mean) positions of the wavepackets during each of the laser-atom interactions. This term is also obtained directly from the atomic trajectories.

The semiclassical path-integral formalism that we used to extract G contains several approximations. An exact calculation would involve integration of Schödinger's equation for a wavepacket subject to the laser-pulse sequence and the gravitational potential of the source mass and Earth. Such a calculation is computationally intractable. We estimated that any deviation from the path-integral formalism resulting from the Gaussian wavepackets and nonquadratic potential terms from the Pb source was negligible (*20*). Terms associated with the Earth rotation were determined to produce negligible phase shifts in the differential phase shift.

Combining the above model with the experimentally measured phase shift, we determined a value of $G = 6.696 \times 10^{-11} \pm 0.037 \times 10^{-11}$ $m^3/(kg \cdot s^2)$ for data run 1. The integration period lasted for 54 data sets. The Pb source mass gravitational phase shifts from the two run cycles are represented in Fig. 4. After a study of potential systematic sources of error on the interferometer phase shift, we performed a repeat measurement. In the latter study we placed the Pb mass 0.635 cm higher than for data run 1, maintaining a 27.940-cm displacement cycle. Furthermore, we reordered the individual Pb disks comprising the source mass. We obtained a value of $G = 6.691 \times 10^{-11} \pm 0.041 \times 10^{-11}$ $m^3/(kg \cdot s^2)$ in the second integration period, lasting 39 data sets. We performed analysis tests similar to those used on the first data run, with the results from runs 1 and 2 agreeing to within statistics. Combining the two measurements results in a statistical value of $G = 6.693 \times 10^{-11} \pm 0.027 \times 10^{-11}$ $m^3/(kg \cdot s^2)$. The bottom plot of Fig. 4 shows our data and its agreement with CODATA. Several G measurements mentioned in the text are listed for comparison.

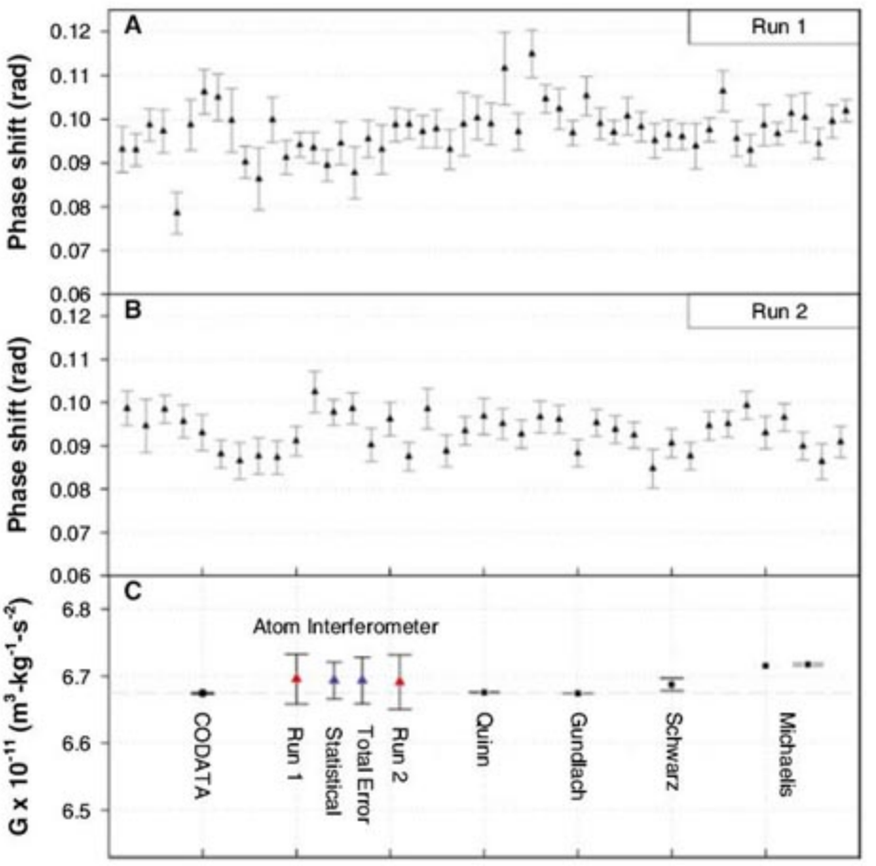


Fig. 4. Data used in the determination of G. (**A**) A data sequence consisting of 51 data sets. Analysis of the phase shifts resulted in a value for $G = 6.696 \times 10^{-11} \pm 0.037 \times 10^{-11}$ $m^3/(kg \cdot s^2)$. (**B**) A second measurement of G with a different initial vertical position of the source mass and a redistribution of the individual discs comprising the Pb stack. The same analysis as the first measurement gave a value of $G = 6.691 \times 10^{-11} \pm 0.041 \times 10^{-11}$ $m^3/(kg \cdot s^2)$. (**C**) Combined results of our measurements agree within statistical uncertainties of each other and of the CODATA value, resulting in $G = 6.693 \times 10^{-11} \pm 0.027 \times 10^{-11}$ $m^3/(kg \cdot s^2)$ (statistical error) and $6.693 \pm 0.021 \times 10^{-11}$ $m^3/(kg \cdot s^2)$ (systematic error). Also shown (left to right) are values reported in Quinn *et al.* (*10*), Gundlach and Merkowitz (*9*), Schwarz *et al.* (*4*), and Michaelis *et al.* (*11*). Error bars show means ± SD.

Knowledge of how the accuracy of our measurement is affected by environmental and device parameters is crucial for a future precision measurement of G with our technique. The results of these tests are summarized in Table 1. Systematic uncertainties limited our experiment to an accuracy of 3 ppt. The dominant systematics were the knowledge of the atom initial position and velocity.

The source mass was specified as 99.99% pure Pb. We characterized the dimensions of the disks and steel plate to 1 ppt with precision calipers and the mass to 40 ppm using a calibrated scale. A volume displacement technique was used to test the density homogeneity of small samples cut from a number of separate Pb disks not used in the G measurement. Density variation was determined to be 260 ppm. Estimating an upper bound of less than 1% radial and longitudinal density inhomogeneities contributes a 0.02% systematic in determining G.

The small residual magnetization of the steel plate used to hold the Pb caused spurious phase shifts through two mechanisms. These shifts were studied by applying additional, much stronger, magnetic sources to the platform. We characterized magnetic fields with the use of Doppler-free Raman spectroscopy of magnetically sensitive hyperfine transitions (an in situ method) and also with the use of a fluxgate magnetometer. The first shift results from second-order Zeeman shifts associated with magnetic field gradients present during the fountain sequence. Propagation reversal of the Raman beams suppresses this shift, but it cannot be completely eliminated because the reversed interferometer paths differ slightly from the nonreversed paths. The second shift results from a Coriolis phase shift that originated in an additional transverse velocity component induced by the presence of a residual magnetic field during the launch. On the basis of the measured residual magnetization of the plate, we estimated that each of these shifts contributed systematic uncertainties of 0.1%.

The gravitational potential of the Pb cylindrical geometry afforded an insensitivity to radial displacements at the 10-ppm level over 5 mm of radial translation. However, we initially found a 5-mrad/mm dependence on the position of the Pb in the plane perpendicular to the interferometer axis. In situ measurements of the magnetic field

at various source mass positions revealed source mass–dependent eddy magnetic fields induced by the switching of the MOT coils. After we implemented a controlled decrease of the field (rather than a sudden switch), we no longer saw a statistically significant dependence on the transverse position of the source for deviations as large as 1 cm.

We individually offset other parameters to values beyond accepted operating characteristics of the gradiometer, often to the point at which the interferometer fringe contrast decreased to below ~10% (typical contrast was 25%). These variables included Doppler sensitive π- and $\pi/2$-pulse lengths, position of atoms in the detection probe beam, detection efficiency, launch angle, off-resonant Raman light, initial $m_f \neq 0$ population (where m_f indicates the Zeeman sublevel), scattering from the background Cs vapor, Raman light intensity, and Raman wavefront quality. Experiments involved the measurement of the Pb-induced phase shift for large offsets in each of the above parameters. At these large offsets, we observed no systematic dependence on the mass displacement signal. From these measurements, we inferred that small drifts of these parameters in time do not contribute systematic offsets in our determination of G.

We looked for systematic effects in our analysis by varying the analysis procedures and parameters. For example, we varied the contrast and outlying phase thresholds used to filter wild points. In this case, we found that for all sets in runs 1 and 2, the inferred values for G agreed to within the statistics. We also studied the effect of the scaling parameters used within the normalized detection scheme to search for possible bias (*25*). Improper normalization resulted in values for G within statistics, whereas the statistical uncertainties increased for large deviations from optimal parameters.

Our demonstrated proof-of-principle measurement of the Newtonian constant of gravity based on atom interferometric measurement of gravity-induced phase shifts presents a technique for the measurement of G not subject to the known and hidden systematics of previous measurements. Since the completion of this work, the experiment of Tino *et al.* (*26*) has begun construction of an atom interferometer apparatus with the goal of increasing the sensitivity and decreasing the systematics to perform a measurement of $\delta G/G = 10^{-4}$.

References and Notes

1. P. J. Mohr, B. N. Taylor, *Rev. Mod. Phys.* **77**, 1 (2005).
2. U. Kleinevoß, H. Meyer, A. Schumacher, S. Hartmann, *Meas. Sci. Technol.* **10**, 492 (1999).
3. F. Nolting, J. Schurr, S. Schlamminger, W. Kündig, *Meas. Sci. Technol.* **10**, 487 (1999).
4. J. P. Schwarz, D. S. Robertson, T. M. Niebauer, J. E. Faller, *Science* **282**, 2230 (1998).
5. C. V. Boys, *Philos. Trans. R. Soc.* **186**, 1 (1895).
6. P. R. Heyl, P. Chrzanowski, *J. Res. Natl. Bur. Std. U.S.* **29**, 1 (1942).
7. G. T. Gillies, *Rep. Prog. Phys.* **60**, 151 (1997).
8. P. J. Mohr, B. N. Taylor, *Rev. Mod. Phys.* **72**, 351 (2000).
9. J. H. Gundlach, S. M. Merkowitz, *Phys. Rev. Lett.* **85**, 2869 (2000).
10. T. J. Quinn, C. C. Speake, S. J. Richman, R. S. Davis, A. Picard, *Phys. Rev. Lett.* **87**, 111101 (2001).
11. W. Michaelis, H. Haars, R. Augustin, *Metrologia* **32**, 267 (1995).
12. K. Kuroda, *Meas. Sci. Technol.* **10**, 435 (1999).
13. W. Michaelis, J. Melcher, H. Haars, *Metrologia* **41**, L29 (2004).
14. A. Peters, K. Y. Chung, S. Chu, *Metrologia* **38**, 25 (2001).
15. J. M. McGuirk, G. T. Foster, J. B. Fixler, M. J. Snadden, M. A. Kasevich, *Phys. Rev. A.* **65**, 033608 (2002).
16. P. Berman, Ed., *Atom Interferometry* (Academic Press, New York, 1996).
17. K. Bongs, R. Launay, M. A. Kasevich, *Appl. Phys. B* **85**, 602 (2006).
18. P. Storey, C. Cohen-Tannoudji, *J. Phys. II* **4**, 1999 (1994).
19. M. J. Snadden, J. M. McGuirk, P. Bouyer, K. G. Haritos, M. A. Kasevich, *Phys. Rev. Lett.* **81**, 971 (1998).
20. Materials and methods are available as supporting material on *Science* Online.
21. E. L. Raab, M. Prentiss, A. Cable, S. Chu, D. E. Pritchard, *Phys. Rev. Lett.* **59**, 2631 (1987).
22. M. Kasevich *et al.*, *Phys. Rev. Lett.* **66**, 2297 (1991).
23. J. M. McGuirk, G. T. Foster, J. B. Fixler, M. A. Kasevich, *Opt. Lett.* **26**, 364 (2001).
24. G. T. Foster, J. B. Fixler, J. M. McGuirk, M. A. Kasevich, *Opt. Lett.* **27**, 951 (2002).
25. Detection-laser powers were periodically adjusted to optimize the detected atom number and signal-to-noise ratio, which results in a slight change of the normalization coefficients, $\alpha_{L,U}$. The normalized $|4,0\rangle$ population is $S_3/(S_3 - \alpha_{L,U}S_2)$, where $S_{3,2}$ is proportional to the $|4,0\rangle$ and $|4,0\rangle - |3,0\rangle$ signal, respectively.
26. A. Bertoldi *et al.*, *Eur. Phys. J. D* **40**, 271 (2006).
27. This work was supported by grants from the ONR and NASA.

Supporting Online Material

www.sciencemag.org/cgi/content/full/315/5808/74/DC1
Materials and Methods
References

22 September 2006; accepted 22 November 2006
10.1126/science.1135459

Conductance-Controlled Point Functionalization of Single-Walled Carbon Nanotubes

Brett R. Goldsmith,[1] John G. Coroneus,[2] Vaikunth R. Khalap,[1] Alexander A. Kane,[1] Gregory A. Weiss,[2,3] Philip G. Collins[1]*

We used covalent attachments to single-walled carbon nanotubes (SWNTs) to fabricate single-molecule electronic devices. The technique does not rely on submicrometer lithography or precision mechanical manipulation, but instead uses circuit conductance to monitor and control covalent attachment to an electrically connected SWNT. Discrete changes in the circuit conductance revealed chemical processes happening in real time and allowed the SWNT sidewalls to be deterministically broken, reformed, and conjugated to target species. By controlling the chemistry through electronically controlled electrochemical potentials, we were able to achieve single chemical attachments. We routinely functionalized pristine, defect-free SWNTs at one, two, or more sites and demonstrated three-terminal devices in which a single attachment controls the electronic response.

Covalently linking a single molecule of interest between two electrical conductors enables the electrical interrogation of that molecule as it dynamically interacts with the surrounding environment. In practice, however, working single-molecule devices remain exceedingly difficult to fabricate (*1*). Successes based on very small electrode gaps fabricated lithographically (*2*), electrically (*3*, *4*), or by scanning probe techniques (*5*, *6*) generally suffer from low fabrication throughput; electrical, mechanical, and chemical instabilities; poorly defined bonding to the molecule of interest; and, sometimes, inconclusive proof that only a single molecule is addressed.

Single-walled carbon nanotubes (SWNTs) have several favorable characteristics for building high-quality, single-molecule devices. Electrically, they are high-conductivity, one-dimensional (1D) conductors that can deliver signals to and from attached molecules. Chemically, SWNTs have long, inert sidewalls but reactive ends to which the tools of organic chemistry can covalently attach a wide variety of species (*7*). Geometrically, SWNTs' small profile maximizes access to the target molecule by reagents, optical probes, or electrostatic fields. Many strategies for building functioning, nanometer-scale circuits have focused on complex manipulation or high-resolution lithographies (*8*–*11*). Guo *et al.*, for example, have beautifully demonstrated single-molecule junctions in broken SWNTs by combining <10-nm lithography with plasma etching (*11*).

Here, we describe an alternative technique that does not require high-resolution lithography and is effective for molecules of any size. The general scheme is to fabricate circuits using individual SWNTs and then use the SWNT conductance G as a real-time indicator of SWNT chemical modification. With the use of electrochemically driven reactions, the introduction of functional groups can be electronically controlled and monitored with microsecond tempo-

[1]Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA. [2]Department of Molecular Biology and Biochemistry, University of California, Irvine, CA 92697, USA. [3]Department of Chemistry, University of California, Irvine, CA 92697, USA.

*To whom correspondence should be addressed. E-mail: collinsp@uci.edu

ral resolution, so that point functionalization can be achieved with better than 90% yield.

Our initial experiments were guided by previous studies on highly oriented pyrolytic graphite crystals (HOPG), as this surface is analogous to the SWNT sidewall. The electrochemical oxidation of HOPG is extensively documented, especially for reactions with strong acids (*12*). At moderate oxidative potentials, these acids add delocalized, positive carriers to HOPG; at higher potentials, however, anions react covalently through addition to the carbon sheet. This bonding disrupts the sp^2 carbon network with sp^3-conjugated carbons, markedly reduces HOPG conductivity, and ultimately leads to the disintegration of HOPG into the insulator graphite oxide (*12*).

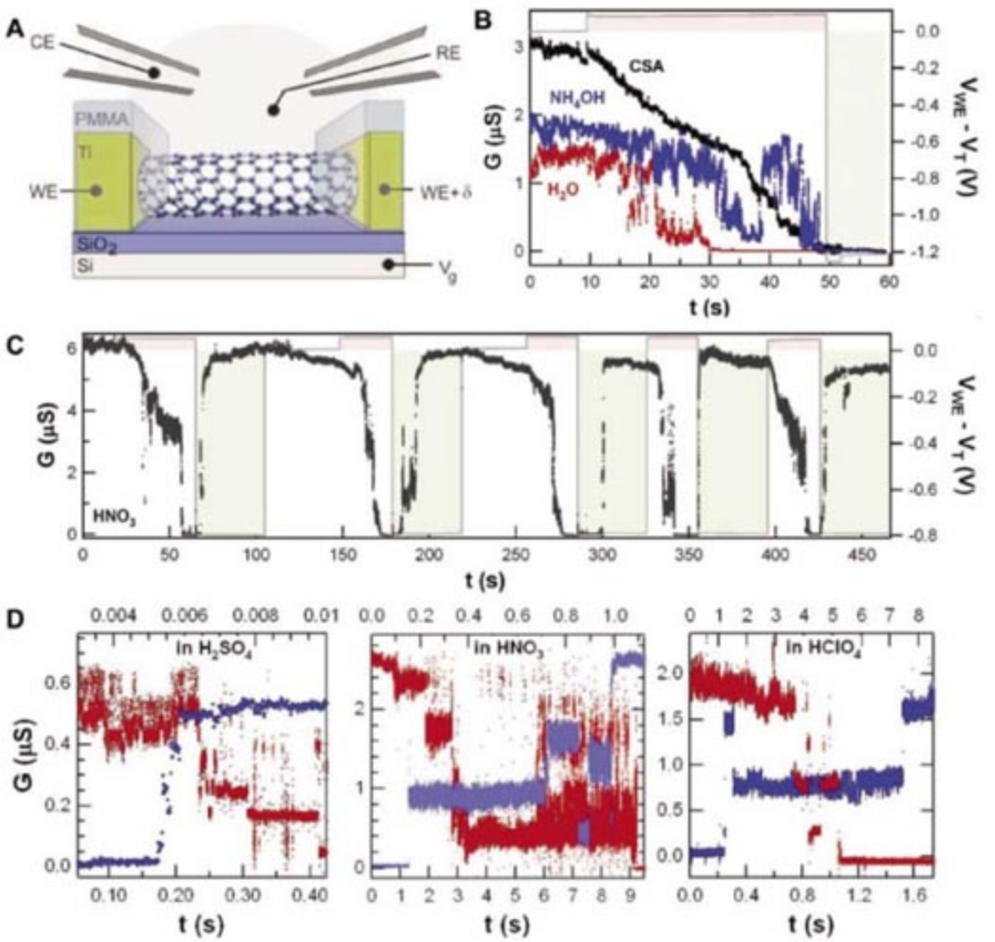


Fig. 1. (**A**) Schematic of SWNT circuit electrochemical modification. The dynamics of the circuit conductance *G* during modification differs between electrolytes. CE, counter electrode; RE, reference electrode; WE, working electrode. (**B**) In most electrolytes, SWNTs oxidize to an open circuit above an electrolyte-dependent threshold V_T (shaded red) and rarely recover under a reducing potential (shaded green). (**C**) In strong acids, reduction results in nearly complete recovery of conductance and the same SWNT may be redox cycled multiple times. (**D**) Higher-resolution oxidation (red, lower time scale) and reduction (blue, upper time scale) traces clarify a fine structure of abrupt jumps among metastable intermediate values. Three different strong acids shown here exhibit similar behaviors. All data in (C) and (D) were acquired at 100 kHz. The reduction portions of *G* in (C) have been scaled up 35% to adjust for the electrostatic gating that occurs at the reducing potential.

Table 1. SWNT electrochemical oxidation characteristics in different electrolytes.

Electrolyte	V_T vs. Pt (V)*	V_T vs. SCE (V)†	ΔV_T (V)	G/G_{init}
HNO_3 (15.3 M)	0.55	1.6	−0.15	80%
$HClO_4$ (7.7 M)	0.70	—	−0.3	90%
H_2SO_4 (18.0 M)	0.80	1.5	−0.1	90%
H_2SO_4 (1.0 M)	0.90	1.4	−0.2	80%
H_3PO_4 (14.7 M)	0.90	—	0	90%
HBr (9.1 M)	0.95	—	0	30%
HCl (12.1 M)	1.10	1.6	0	60%
H_2O (deionized)	1.25	1.5	0	60%
$KMnO_4$ (6 mM)	0.70	—	—	None
CSA‡ (0.2 M)	1.25	—	—	None
NH_4OH (4.3 M)	1.00	1.4	—	None
NaOH (2.6 M)	0.70	1.5	—	None

*All errors are +0.10/−0.05 V except for H_3PO_4 (±0.2). †All errors are ±0.1. ‡Anhydrous camphorsulfonic acid, in acetonitrile.

Qualitatively similar effects have been reported for bulk SWNTs (*13*, *14*) but not, to our knowledge, on single, isolated SWNTs, perhaps because of the characterization challenges involved in observing reactions with single- or few-bond resolution (*15*). However, as shown here, the SWNT's low-bias conductance *G* is sufficiently sensitive to allow detection. Multiple independent theoretical models have recently calculated *G* in the presence of single chemical defects, vacancies, or adducts, and have predicted changes (ΔG) on the order of 50% (*16*–*21*). The remarkable sensitivity of *G* to small compositional changes arises from the 1D conductance of SWNTs, which amplifies the effect that a single sp^3-hybridized carbon can have among 10^6 sp^2-hybridized carbons. Below, we demonstrate the tremendous sensitivity of *G* and the potential identification of single-bond chemical events, and describe the use of different functionalization schemes to selectively tailor the resulting sites.

The fabrication and electrochemical modification of samples is discussed in (*22*) and follows our previous reports (*23*, *24*). Briefly, SWNTs in a narrow diameter range of 1.0 to 1.2 nm are grown by catalyst-assisted chemical vapor deposition (CVD) on thermally oxidized Si wafers. After growth, the SWNTs are contacted by multiple metal electrodes fabricated by optical lithography, and additional lithography is optionally defined on top of the devices to protect the metal and metal-SWNT interfaces with insulating coatings. To prepare them for controlled oxidation, completed devices are mounted in an electrochemical cell (*23*–*26*) and contacted by movable probe tips or fixed wirebonds. Platinum counter and reference microelectrodes control the liquid potential (Fig. 1A), although the harsh oxidation conditions require independent calibration of the Pt pseudo-reference against a standard calomel electrode (SCE) in each electrolyte.

The chemical modifications and attachments described below have been tested using Au, Pd, Ti, and TiN electrode metals with and without protective polymethyl methacrylate (PMMA) coatings over the metal. We have determined that the electrode metal and the metal-SWNT interface do not affect the results presented here (*22*), so we restrict our attention to Ti electrodes unless otherwise noted, as these devices are the least complicated to prepare. The native oxide of the Ti contact metal is sufficient to suppress unwanted electrochemical currents without additional PMMA insulation or lithography (fig. S3) (*27*).

The electrochemical modification of a SWNT is electrolyte-dependent, and Table 1 summarizes the electrolytes tested for this report. In every electrolyte, G is stable across a range of electrochemical potentials, but above an oxidative threshold V_T, the value of G decreases to zero (Fig. 1). Oxidation in bases and weak acids tends to be electrochemically irreversible; Fig. 1B depicts typical results in H_2O, HN_4OH, and nonaqueous camphorsulfonic acid (CSA). On the other hand, the strong acids H_2SO_4, HNO_3, $HClO_4$, and H_3PO_4 exhibit a very different behavior in which the application of a reductive potential after oxidation reverses the drop in G (*23*). This recovery effectively constitutes a redox cycle of the SWNT sidewall and allows the same SWNT to be repeatedly cycled (Fig. 1C). The average degree of recovery is characterized in Table 1 as a percentage of the pristine SWNT's initial G_{init}, but in general the strong acids all behave similarly. We find metallic and semiconducting SWNTs to have indistinguishable redox behaviors, perhaps because the liquid gate is very effective at moving the Fermi level E_F into the valence band at the large oxidative potentials used here (*25*, *26*). We also note that variability in V_T is primarily caused by chemical instabilities of the Pt microreference, so that within the experimental error it is a constant $V_T = 1.6 \pm 0.1$ V versus SCE. This value may represent the true oxidation threshold of defect- and edge-free graphene, and as such it slightly exceeds what can be achieved with high-quality HOPG (*28*, *29*).

In both reversible and irreversible electrolytes, the dynamics of the SWNT failure are revealing. Precisely biased just below the threshold V_T, G decreases continuously in time before stabilizing near $0.7G_{init}$. More typically, coarse potentiostat control will exceed V_T, and in this case G continues its decrease through a complex sequence of abrupt, discrete steps. These steps are most reliably observed in the strong acids (Fig. 1, C and D). During both oxidation (red) and subsequent reduction (blue), $G(t)$ repeatedly samples and stabilizes at discrete conductance values, then jumps among these values in a period equal to or less than the temporal resolution of the data (10 μs). The intermediate terraces during oxidation exhibit excellent alignment with those during reduction, with little indication of anion dependence (*19*). Although such good alignment is not observed in every measurement, the reduction and oxidation traces frequently have similar numbers of terraces and step heights (ΔG). The different time scales used for each trace in Fig. 1D indicate the degree to which the chemical processes are sensitive to the applied voltages once a reaction threshold is exceeded.

In either the oxidized or reduced state, no structural damage was observed on these SWNTs within the resolution limits of atomic force microscopy (AFM), and this finding, combined with the redox reversibility, rules out the possibility that carbon atoms are being permanently removed from the SWNT. Instead, we interpret the changes in G to be indicative of the same chemistries that occur on HOPG, except amplified by the 1D nature of a SWNT. The initial, continuous changes in G can be attributed to charge transfer–induced disorder and the discrete steps to covalent oxidation (*12*–*14*). The short time required for one step ΔG to occur and the long latencies between steps suggest a stochastic process consistent with individual oxidation events; the step heights observed are consistent with the theoretical predictions for such events (*16*–*21*). We conclude that each step likely corresponds to the formation of a single covalent C-O bond between the SWNT and the nitrate, sulfate, or other conjugate base, in the same ways that such anions form adducts with the HOPG basal plane (*12*). The detailed dynamics and metastability in G then likely reflect the different leaving-group abilities for the elimination reaction.

Despite the apparent reversibility of the cycling, multiple redox cycles on the same SWNT result in both random and gradual degradation, even in the best electrolytes. For example, Fig. 1C shows a 20% degradation after five redox cycles in HNO_3. We have extensively studied SWNTs before and after single redox cycles and have identified the cause of this degradation: Residual SWNT modifications remain even when G recovers to $>0.9G_{init}$. The first indication of this residual damage is a small, one-time threshold reduction ΔV_T that occurs between the first and all subsequent redox cycles. This shift indicates that subsequent oxidation can take advantage of some residual disorder and/or enhanced reactivity.

Second, we observe a substantial change in the electronic behavior of every metallic SWNT measured. Normally, the G of a metallic SWNT is insensitive to a nearby gate electrode biased at voltage V_g. As shown in Fig. 2, A and B, single redox cycles electronically alter metallic SWNTs and make them gate-sensitive. Experimentally and theoretically, it has been established (*30*–*32*) that a point defect embedded in a metallic SWNT can cause this V_g modulation.

Third, we have used scanning probe techniques to investigate SWNTs before and after redox cycling. Although the residual disorder is topologically invisible, scanning gate microscopy (SGM) and Kelvin force microscopy (KFM) can resolve it (*33*). SGM is a technique in which V_g is locally applied with an AFM cantilever to a small region of a SWNT. The resulting map of G (V_g) spatially identifies any SWNT regions that contribute to gate sensitivity. Figure 2C is an AFM image of a metallic SWNT and, in false color, its SGM map showing a V_g-sensitive site added by redox cycling. We observe the introduction of similar sites to semiconducting SWNTs, although in this case the change in $G(V_g)$ is naturally smaller than in Fig. 2A. Finally, we observe that annealing the redox-cycled samples (400°C in vacuum or N_2) fully restores both the initial $G(V_g)$ and the oxidation threshold V_T.

We conclude that electrochemical reduction does not return a SWNT to its pristine chemical state, despite the near-perfect recovery in G. Instead, reduction further alters the SWNT's chemical state at one or more sites. Because of the limited active area, we have been unable to directly detect the resulting chemical functionalities by common optical spectroscopies, but identical acid oxidation produces adducts such as hydroxides, epoxides, and ethers on HOPG and bulk SWNTs (*12*–*14*). Of these, divalent ethers are a particularly attractive candidate for the redox-cycled SWNTs because they have been predicted to minimally affect G (*21*, *34*, *35*). Whereas other functionalities produce sp^3 hybridizations, the ethers have sp^2 carbons that only weakly scatter free carriers. Figure 2D outlines a scheme adopted from HOPG in which a strong acid oxidizes the SWNT sidewall and, upon subsequent reduction, leaves behind the ether functionalization.

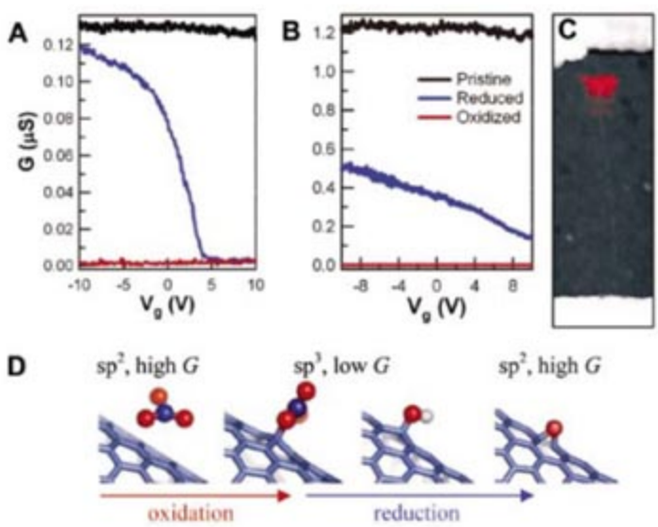


Fig. 2. (**A** and **B**) After a single redox cycle, metallic SWNTs show strong V_g dependences. $G(V_g)$ is flat before oxidation (black), zero after oxidation (red), and then semiconductor-like after reduction (blue). Subsequent cycles have lesser effects on G, but flat $G(V_g)$ curves were never recovered. It is also common for the G recovery to be incomplete in metallic SWNTs (B). (**C**) A composite of AFM topography in grayscale and SGM in red identifies a local region responsible for the gate sensitivity of the device in (A). (**D**) Proposed chemical process for HNO_3 oxidation and reduction, in which the redox cycle is not fully reversible but instead leaves behind a C-OH or C-O-C residue. The latter has sp^2 conjugation and restores the SWNT conductivity.

Regardless of the exact chemical pathway, the residual sites provide a chemical handle for further sidewall functionalization. Additional reactions can be promoted on either the oxidized or reduced SWNTs. However, the stochastic nature of the original damage means that these reactions will tailor a very limited number of functionally active sites. Thanks to the sensitivity of G, electronic triggering can give ready control over the degree of covalent oxidation, and the electrolyte used can provide flexible control over the functional adducts.

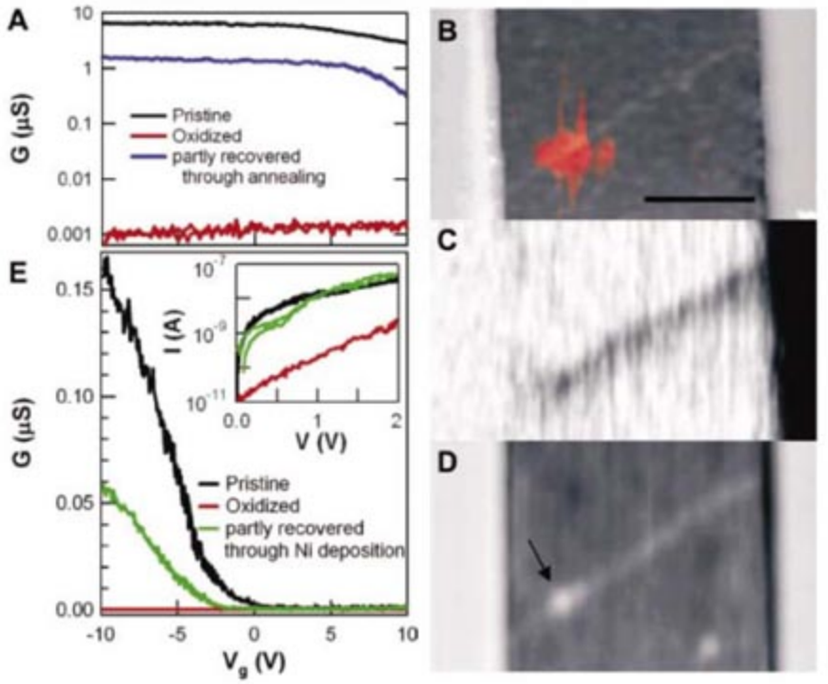


Fig. 3. (**A**) $G(V_g)$ for a quasi-metallic SWNT electrochemically oxidized and exposed to $KMnO_4$ (red). The additional $KMnO_4$ oxidation permanently damages the SWNT, so that even thermal annealing does not fully recover G (blue). (**B**) A composite of AFM topography (grayscale) and SGM (red) identifies the region responsible for the enhanced gate sensitivity of an oxidized SWNT. (**C**) KFM imaging shows that most of the potential applied across the SWNT is dropped across this gate-sensitive region. (**D**) Selective electrodeposition of Ni identifies this site as more chemically reactive than the rest of the SWNT. Scale bar, 500 nm. (**E**) A similar experiment on a semiconducting SWNT in which the device is open-circuited before Ni deposition and conducting afterward (green). $G(V_g)$ reflects the SWNT band structure as well as the extra resistance associated with the Ni-coated, oxidized region. Inset shows corresponding current-voltage characteristics.

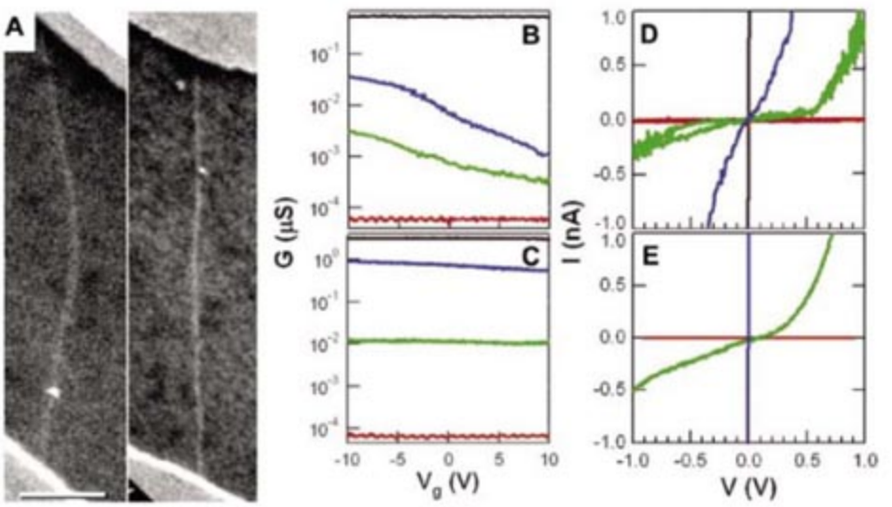


Fig. 4. (**A**) Scanning electron micrographs of SWNT devices incubated with streptavidin after controlled oxidation. Streptavidin covalently binds to each oxidation site, which has been activated by treatment with EDC and NHS. Nonspecific adsorption is minimized by a polysorbate surfactant. Each SWNT is visible because it is a conductor sitting on an insulating surface; the proteins are not visible but are labeled with 20-nm Au particles for imaging. Scale bar, 500 nm. (**B** and **C**) $G(V_g)$ curves at each stage in the chemical treatment show the pristine (black curves), oxidized (red), and protein-conjugated (green) chemical states, with a substantial recovery in conductance upon successful conjugation. (**D** and **E**) The corresponding current-voltage curves show decidedly nonlinear behaviors. Note that the high-conductance states (black, blue) and the oxidized, low-conductance state (red) overlap the graphs' axes. Thermal annealing of the samples does not fully recover any of the conjugated SWNT devices (blue curves).

For example, further oxidation of SWNTs has been tested in aqueous $KMnO_4$ (6.5 mM). Pristine SWNT devices are unaffected when exposed to $KMnO_4$. Furthermore, performing a single oxidation-reduction cycle in acid is usually insufficient to make a SWNT reactive with $KMnO_4$. Alternatively, SWNTs that have been oxidized but not reduced are irreversibly changed by exposure to $KMnO_4$ for 30 s. These devices do not conduct at low bias, even after electrochemical reduction, and thermal annealing only partially recovers the device characteristics. Careful examination of $G(t)$ traces acquired before, during, and after $KMnO_4$ application can prevent accidental reduction before $KMnO_4$ reaction. Figure 3A shows a typical $G(V_g)$ characteristic of a metallic SWNT before modification, after acid and $KMnO_4$ oxidation, and then after thermal annealing. The new gate sensitivity in the device can be mapped by SGM to a local region of the SWNT sidewall (Fig. 3B). The complementary technique of KFM, in which the electrostatic potential of the SWNT is mapped, shows a potential drop in the same region (Fig. 3C).

We have also performed selective electrochemical deposition, a technique specially designed to label point defects on the basis of their enhanced reactivity (*24*). Pristine SWNTs found to have no electrochemically active sites were routinely labeled by selective Ni deposition after a redox cycle, and this labeling usually occurred at a single site (*22*). More remarkably, Ni could be deposited on open-circuit SWNTs that had been oxidized with $KMnO_4$. In this case, the Ni deposit is not only visible (Fig. 3D) but also reconnects the electrical circuit. Figure 3E shows $G(V_g)$ electrical characterization of a semiconducting SWNT before and after oxidation and then after Ni deposition. The oxidation decreases G by a factor of 1000, producing a characteristic behavior similar to gap tunneling (*23*). Nickel deposition closes this tunneling gap (but does not affect the SWNTs' semiconducting band structure) to reproduce the initial $G(V_g)$ behavior but with a higher on-state resistance. Kept in air, these devices become insulating again as the Ni deposition oxidizes, but storage in vacuum stabilizes the electrical characteristics indefinitely.

The electrochemical and scanning probe techniques have resolution limits of 5 and 50 nm, respectively, and cannot distinguish between single- and multiple-bond disruptions to the SWNT, but they do rule out randomly distributed or widely spaced damage. Furthermore, the presence of an oxygen-containing functionality is virtually certain, given that Ni deposition proceeds with nearly 100% yield. Reactions with higher chemoselectivity are required to achieve

higher resolution and further discriminate among possible chemical modifications.

For example, the highly chemoselective activation of carboxyl groups by treatment with *N*-ethyl-*N*′-(3-dimethyl aminopropyl) carbodiimide (EDC) and *N*-hydroxysuccinimide (NHS) is used to attach proteins to carboxy-terminated surfaces via amide linkages (*36*). In a third series of chemical experiments, redox-cycled SWNTs were treated with EDC and NHS in an attempt to covalently link Au-labeled streptavidin to SWNT carboxyls (*37*). Only one attachment was observed on 10 redox-cycled SWNT devices, which shows that the acid treatment does not generally produce sidewall carboxy groups. This result is in agreement with our speculation that ethers are the dominant residual functionality. However, four of nine devices showed protein attachments when additional $KMnO_4$ oxidation was added to the procedure. $KMnO_4$ readily converts hydroxides to carboxy groups, so we can conclude with some certainty that the electronic maps in Fig. 3, B and C, are related to carboxy-functionalized sites.

Figure 4A and fig. S5A show scanning electron micrographs of successful protein attachments. In each image, a single streptavidin-coated Au particle is integrated into a functioning SWNT device. Electrical characterization of these devices shows that each circuit is insulating immediately before streptavidin attachment but conducting afterward (Fig. 4, B to E). G is critically dependent on the presence of the protein, just as in the case of the Ni decorations, even though the current does not flow through the entire protein or gold particle per se. As in Fig. 3E, the measured G reflects the series combination of SWNT band structure and protein-linked oxidation gaps, and these devices could exhibit strong biosensitivity, as previously shown for noncovalently coated SWNTs (*38*).

The usefulness of G to monitor and control chemistry in situ and in real time arises from its sensitivity to a few single-bond redox events. Paradoxically, SWNT devices never drop directly to $G = 0$ without intermediate steps and terraces, and questions remain regarding the exact nature of these terraces. Additional experiments may distinguish among intermediate chemical states, oxidation cascades among neighboring carbon atoms, or other possible mechanisms.

References and Notes

1. G. Maruccio *et al.*, *Electroanalysis* **16**, 1853 (2004).
2. A. A. Tseng, A. Notargiacomo, T. P. Chen, *J. Vac. Sci. Technol. B* **23**, 877 (2005).
3. H. Park, A. K. L. Lim, A. P. Alivisatos, J. Park, P. L. McEuen, *Appl. Phys. Lett.* **75**, 301 (1999).
4. D. R. Strachan *et al.*, *Appl. Phys. Lett.* **86**, 43109 (2005).
5. W. Ho, *J. Chem. Phys.* **117**, 11033 (2002).
6. N. J. Tao, *J. Mater. Chem.* **15**, 3260 (2005).
7. S. Banerjee, T. Hemraj-Benny, S. S. Wong, *Adv. Mater.* **17**, 17 (2005).
8. P. W. Chiu, G. S. Duesberg, U. Dettlaff-Weglikowska, S. Roth, *Appl. Phys. Lett.* **80**, 3811 (2002).
9. Y. X. Zhou, A. T. Johnson, J. Hone, W. F. Smith, *Nano Lett.* **3**, 1371 (2003).
10. M. R. Diehl *et al.*, *ChemPhysChem* **4**, 1335 (2003).
11. X. Guo *et al.*, *Science* **311**, 356 (2006).
12. K. Kinoshita, *Carbon—Electrochemical and Physicochemical Properties* (Wiley Interscience, New York, 1988).
13. G. U. Sumanasekera *et al.*, *J. Phys. Chem. B* **103**, 4292 (1999).
14. R. Graupner *et al.*, *Phys. Chem. Chem. Phys.* **5**, 5472 (2003).
15. M. Burghard, *Surf. Sci. Rep.* **58**, 1 (2005).
16. X. Lu, Z. F. Chen, *Chem. Rev.* **105**, 3643 (2005).
17. J. Zhao *et al.*, *ChemPhysChem* **6**, 598 (2005).
18. Y. W. Son, J. Ihm, M. L. Cohen, S. G. Louie, H. J. Choi, *Phys. Rev. Lett.* **95**, 216602 (2005).
19. Y. S. Lee, M. B. Nardelli, N. Marzari, *Phys. Rev. Lett.* **95**, 076804 (2005).
20. J. Eom *et al.*, *Physica B* **376–377**, 7 (2006).
21. Y.-S. Lee, N. Marzari, *Phys. Rev. Lett.* **97**, 116801 (2006).
22. See supporting material on *Science* Online.
23. J. Mannik, B. R. Goldsmith, A. Kane, P. G. Collins, *Phys. Rev. Lett.* **97**, 016601 (2006).
24. Y. Fan, B. R. Goldsmith, P. G. Collins, *Nat. Mater.* **4**, 906 (2005).
25. M. Krüger, M. R. Buitelaar, T. Nussbaumer, C. Schonenberger, L. Forro, *Appl. Phys. Lett.* **78**, 1291 (2001).
26. S. Rosenblatt *et al.*, *Nano Lett.* **2**, 869 (2002).
27. I. Heller *et al.*, *Nano Lett.* **5**, 137 (2005).
28. J. O. Besenhard, H. Mohwald, J. J. Nickl, *Synth. Met.* **3**, 187 (1981).
29. A. Metrot, J. E. Fischer, *Synth. Met.* **3**, 201 (1981).
30. M. Bockrath *et al.*, *Science* **291**, 283 (2001).
31. S. J. Tans, C. Dekker, *Nature* **404**, 834 (2000).
32. M. Freitag, A. T. Johnson, S. V. Kalinin, D. A. Bonnell, *Phys. Rev. Lett.* **89**, 216801 (2002).
33. A. Bachtold *et al.*, *Phys. Rev. Lett.* **84**, 6082 (2000).
34. H. Park, J. J. Zhao, J. P. Lu, *Nanotechnology* **16**, 635 (2005).
35. H. Pan, Y. P. Feng, J. Y. Lin, *Phys. Rev. B* **70**, 245425 (2004).
36. J. V. Staros, R. W. Wright, D. M. Swingle, *Anal. Biochem.* **156**, 220 (1986).
37. S. S. Wong, E. Joselevich, A. T. Woolley, C. L. Cheung, C. M. Lieber, *Nature* **394**, 52 (1998).
38. A. Star, J. C. P. Gabriel, K. Bradley, G. Gruner, *Nano Lett.* **3**, 459 (2003).

39. We thank J. Mannik and F. Alim for experimental assistance, and our team collaborators R. Penner and N. Allbritton for helpful conversations. Supported by NSF NIRT grant EF-0404057, NSF CAREER grant DMR-023-9842 (P.G.C.), American Chemical Society Petroleum Research Fund grant 39672-G5M (P.G.C.), National Institute of General Medical Sciences grant R01 GM078528-01 (G.A.W.), and National Institute of Allergy and Infectious Diseases grant R43 AI58365-01 (G.A.W.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/315/5808/77/DC1
Materials and Methods
SOM Text
Figs. S1 to S5

19 September 2006; accepted 7 November 2006
10.1126/science.1135303

Counting Low–Copy Number Proteins in a Single Cell

Bo Huang,[1]* Hongkai Wu,[1]† Devaki Bhaya,[2] Arthur Grossman,[2] Sebastien Granier,[3] Brian K. Kobilka,[3] Richard N. Zare[1]‡

We have designed a microfluidic device in which we can manipulate, lyse, label, separate, and quantify the protein contents of a single cell using single-molecule fluorescence counting. Generic labeling of proteins is achieved through fluorescent-antibody binding. The use of cylindrical optics enables high-efficiency (≈60%) counting of molecules in micrometer-sized channels. We used this microfluidic device to quantify β_2 adrenergic receptors expressed in insect cells (SF9). We also analyzed phycobiliprotein contents in individual cyanobacterial cells (*Synechococcus* sp. PCC 7942) and observed marked differences in the levels of specific complexes in cell populations that were grown under nitrogen-depleted conditions.

Single-cell analysis has become a highly attractive tool for investigating cellular contents (*1*). Unlike conventional methods that are performed with large cell populations, this technology avoids the loss of information associated with ensemble averaging. Recent studies have described methods that can quantify specific proteins inside a single cell (*2*–*4*) by means of integrated fluorescence (including confocal microscopy, flow cytometry, and monitoring fluorescent enzymatic products) and, in another instance (*5*), by single-molecule imaging. These techniques restrict analysis to one or perhaps a few species at a time because of the need to resolve fluorescence from different probes. Moreover, their applications are limited in the cases where the cell environment changes the fluorescence of the reporter molecule (e.g., through quenching or resonance energy transfer) or where endogenous fluorescence interferes with the measurements.

We present a different approach based on manipulating, capturing, and lysing a single cell, followed by chemical separation and analysis of the lysate. This approach sacrifices the possibility of monitoring live cells but gains the ability to quantify multiple targets that cannot be distinguished by fluorescence properties alone. Recent achievements—including the analysis of DNA (*6*), amino acid profiles (*7*), and protein fingerprints (*8*)—use either a capillary or a mi-

[1]Department of Chemistry, Stanford University, Stanford, CA 94305-5080, USA. [2]Department of Plant Biology, Carnegie Institution, Stanford, CA 94305, USA. [3]Department of Molecular and Cellular Physiology and Medicine, Stanford University, Stanford, CA 94305-5345, USA.

*Present address: Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA.
†Present address: Department of Chemistry, Tsinghua University, Beijing 100084, China.
‡To whom correspondence should be addressed. E-mail: zare@stanford.edu

crofluidic platform. Compared with the former platform, microfluidics allows sophisticated cell manipulation and provides the potential for extremely high detection sensitivity as well as high-throughput screening.

Low–copy-number proteins (present at fewer than 1000 molecules per cell) play an important role in cell functioning, including signaling and the regulation of gene expression. Without amplification procedures, their abundance is far below the sensitivity limits of conventional protein analysis methods, such as enzyme-linked immunosorbent assay and mass spectroscopy. To solve this problem, we used single-molecule fluorescence detection, which has been applied to counting DNA or protein molecules in sheathed flows, capillaries, and microfluidic channels (*9*–*11*).

Our single-cell analysis chip for mammalian and insect cells (Fig. 1A) is made of polydimethylsiloxane (PDMS) and has three sections: cell manipulation, electrophoretic separation, and single-molecule counting. For cell manipulation, we used a three-state valve design that has been demonstrated previously (*7*). After the cell was captured in the reaction chamber formed between the three-state valve and a conventional two-state valve, a lysing/labeling buffer was injected into the chamber so that the cell contents were released (Fig. 1B). In our experiment with proteins that are not naturally fluorescent, we used fluorescently labeled antibodies as a generic method to tag target proteins. A nonionic detergent, 1 weight percent (wt%) *n*-dodecyl-β-D-maltoside (DDM), was used as the lysing reagent, which preserved the activity of the antibodies.

After the labeling reaction occurred, the excess labeling reagent was electrophoretically separated from the target proteins. To reduce the transportation loss of the sample, the separation buffer contained 0.1 wt% DDM so that protein adsorption on the hydrophobic PDMS channel walls was essentially suppressed (*12*). A low concentration (~0.005 to ~0.05 wt%) of sodium dodecylsulfate (SDS) was added to generate sufficient electro-osmotic flow by its adsorption to the PDMS surface (*13*). Although we used DDM/SDS separation for all of our experiments, we could have used any other electrophoretic/chromatographic separation method that is compatible with a PDMS microfluidic chip.

At the end of the separation channel, fluorescent molecules were counted by monitoring the number of fluorescence bursts generated when the molecules flowed through a small detection volume. For single-molecule counting, the most common approach to obtaining a high signal-to-noise ratio is to use confocal fluorescence microscopy, but its detection cross section (approximately 500 nm wide by 1 μm high) is much smaller than the cross section of an ordinary microfluidic channel (100 μm by 10 μm), which results in extremely poor detection efficiency (*14*). Several groups have attempted to solve this problem by decreasing the dimensions of the channel or capillary to the nanometer range so that the entire cross section fits into the focus of the confocal microscope (*15*–*17*). Such a small channel dimension, however, could affect the electrophoretic separation of molecules in cell

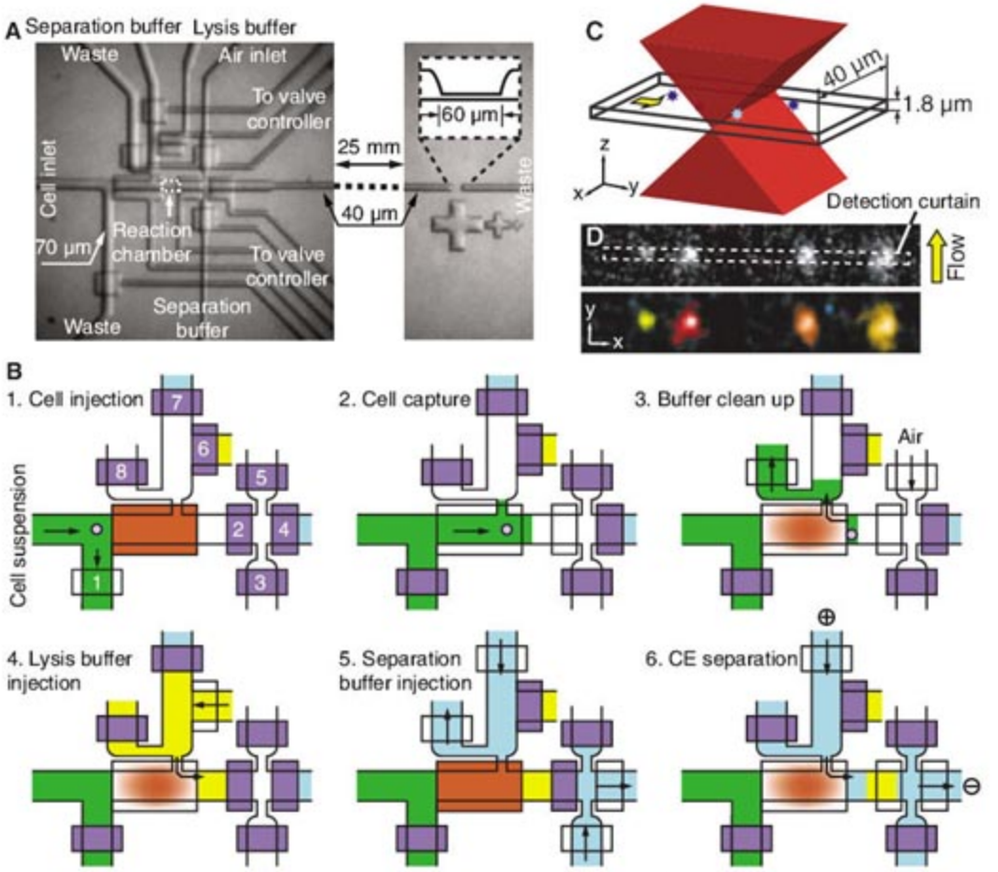


Fig. 1. The single-cell analysis chip. (**A**) Layout of the single-cell chip, showing the cell-manipulation section on the left and the molecule-counting section on the right. (**B**) Analysis procedure for a mammalian or insect cell. (**C**) Schematic illustration of the excitation laser focused by the microscope objective and the dimensions of the molecule-counting channel. (**D**) One frame from the CCD images of fluorescent molecules flowing across the molecule-counting section (upper panel) and the identification results (lower panel).

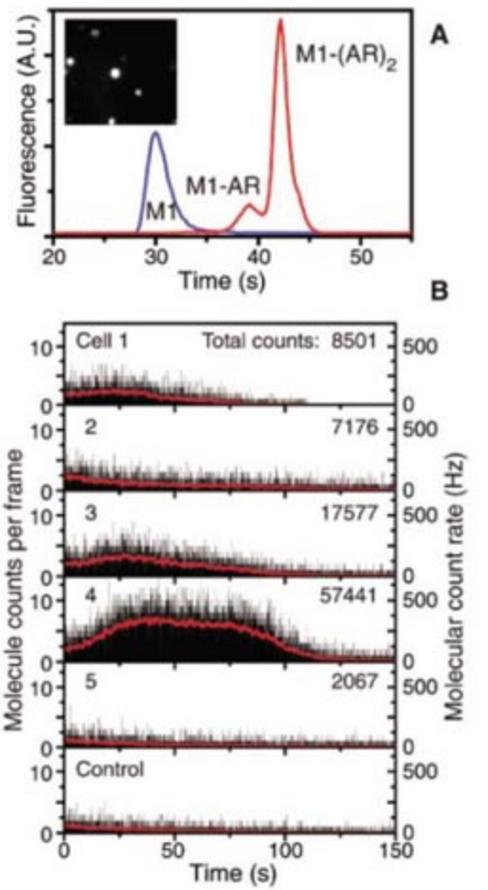


Fig. 2. Analysis of β_2AR in SF9 cells. (**A**) Electropherogram of Cy5-labeled M1 antibody against FLAG (M1); measurements are shown before and after adding an excess amount of purified β_2AR (AR) in a double-T chip. The inset is an image of SF9 cells expressing β_2AR and fluorescently stained with M1. A.U., arbitrary units. (**B**) Molecule-counting results of SF9 cells expressing β_2AR, showing the electropherogram of the M1-AR complexes. The red line represents the average count rate. The total counts are corrected for the counting efficiency, the existence of M1-$(AR)_2$ complexes, and blank counts in the control experiment (5613 molecules). The control experiment was performed by separating lysis/labeling buffer in the reaction chamber with no cells captured. The background counts were mainly caused by the tailing of the large free-antibody peak that came earlier in the electropherogram (not shown).

lysates (*18*, *19*) and could also lead to a clogging of the nanochannel with cell debris.

To resolve the counting-efficiency problem, we widened the excitation laser focus in one direction by using cylindrical optics. The excitation laser beam was focused by a cylindrical lens to form a line at the back focal plane of a high-numerical-aperture objective (fig. S1). When the laser beam emerged from the objective, it was collimated in the direction perpendicular to the channel length so that it illuminated a channel width of tens of micrometers (Fig. 1C). In the other direction, the laser was still tightly focused by the objective to minimize the fluorescence background from out-of-focus excitation. The limit of the channel height by the excitation focus was ~2 μm (fig. S1B), which was slightly relaxed as compared to confocal detection because the laser was less tightly focused. This height was large enough to avoid clogging in our experiments.

With this optical configuration, the excitation laser formed a rectangular, curtain-shaped detection region across the channel. The fluorescence from molecules that passed through the curtain was recorded by an intensified charge-coupled device (CCD) camera. We used a threshold criterion to identify each molecule as a bright spot in an image frame (Fig. 1D). To enhance the fluorescence signal for molecule counting, we slowed down the flow rate when an analyte was expected to pass through the detection curtain.

Because of the variation in excitation laser intensity at different positions in the channel cross section, molecules that passed through the periphery of the channel produced lower fluorescence signals, which could be lost in the background noise. As a result, the detection efficiency varied slightly according to the brightness of a specific sample molecule. We have measured the molecular-counting efficiency of Alexa Fluor 647–labeled streptavidin in a standard double-T chip and found that 60% of the injected molecules are counted (*20*). A general method was then developed to estimate detection efficiencies directly from counting experiments by varying the threshold of molecule identification (*20*). In this way, the actual number of molecules can be derived without additional calibration.

Using the single-cell analysis chip, we quantified the copy number of a human transmembrane protein, β_2 adrenergic receptor (β_2AR), expressed in an insect cell line (SF9). β_2AR is not naturally fluorescent. Therefore, we genetically added a short peptide sequence, the FLAG tag, to the N terminus (*21*), so that it could bind Cy5-labeled monoclonal M1 antibody against FLAG (Cy5-M1) with high affinity (dissociation constant = 2.4 nM) (*22*). The antibody concentration was 20 nM in the lysis/labeling buffer, which caused ~90% of the β_2AR to be fluorescently tagged. β_2AR in SF9 cells can be efficiently extracted with 1 wt% DDM, which is demonstrated by the rapid disappearance of fluorescence from Cy5-M1–stained cells when the lysing buffer is added.

The immunocomplexes of Cy5-M1 and purified β_2AR could be separated with the DDM/SDS system (Fig. 2A). Figure 2B lists the molecule-counting results of five transfected SF9 cells. These cells contained an average of ~1.8×10^4 β_2AR per cell, which was in agreement with the ensemble measurement using cell lysate [4×10^4 β_2AR per cell, detectable by Cy5-M1 binding (*20*)]. The copy numbers of β_2AR in these cells varied between ~2000 and ~60,000. By staining with Cy5-M1 (Fig. 2A, inset), we could observe similar variations in whole-cell fluorescence intensity, although cell staining detects only the β_2AR that is present on the cell surface. This cell-to-cell variation is probably caused by the stochastic nature of virus infection; i.e., individual cells are not infected simultaneously and may contain different numbers of viral particles, resulting in different expression levels of β_2AR.

To demonstrate a biological application of single-cell analysis, we studied the response of the unicellular cyanobacterium *Synechococcus* sp. PCC 7942 (*Synechococcus* hereafter) to the depletion of nitrogen-containing nutrients in the culture medium (chlorosis). Cyanobacteria and some eukaryotic algae use the phycobilisome (PBS), a peripheral membrane protein-chromophore complex, to collect the light energy and transfer it to the photosynthetic reaction centers (*23*). In *Synechococcus* cells, the PBS comprises mainly two pigmented phycobiliproteins (PBPs): phycocyanin (PC), which exists in the peripheral rods, and allophycocyanin (APC), which forms the core structure. It also contains various linker polypeptides that function to hold together the assembly and to tune the complex for efficient energy flow into the photosynthetic reaction centers. A chromophore-containing linker polypeptide designated L_{CM} attaches the PBS to photosystem II on the thylakoid membrane (*23*, *24*).

When grown under conditions in which certain macronutrients (such as nitrogen) are depleted, these cyanobacteria begin to degrade their PBS in an ordered way (first PC, then APC). This process reduces the absorption of excess light energy and provides cells with nutrients from the degraded PBP, helping them to attain a quiescent state almost devoid of the PBS (Fig. 3A, inset) (*23*, *25*). To characterize this process at the ensemble level, we lysed *Synechococcus* that was cultured in nitrogen-replete (+N) or nitrogen-depleted (–N) media using the combination of lysozyme and a nonionic surfactant. The extracted PBP complexes in the lysate were electrophoretically separated on a microfluidic chip. As shown in Fig. 3A, after *Synechococcus* was cultured in a –N medium for more than 72 hours, all peaks related to PC (peaks 1 and 4 to 9) nearly disappeared. The two major PBP peaks that remained correspond to two APC subassemblies in the PBS core (peak 2

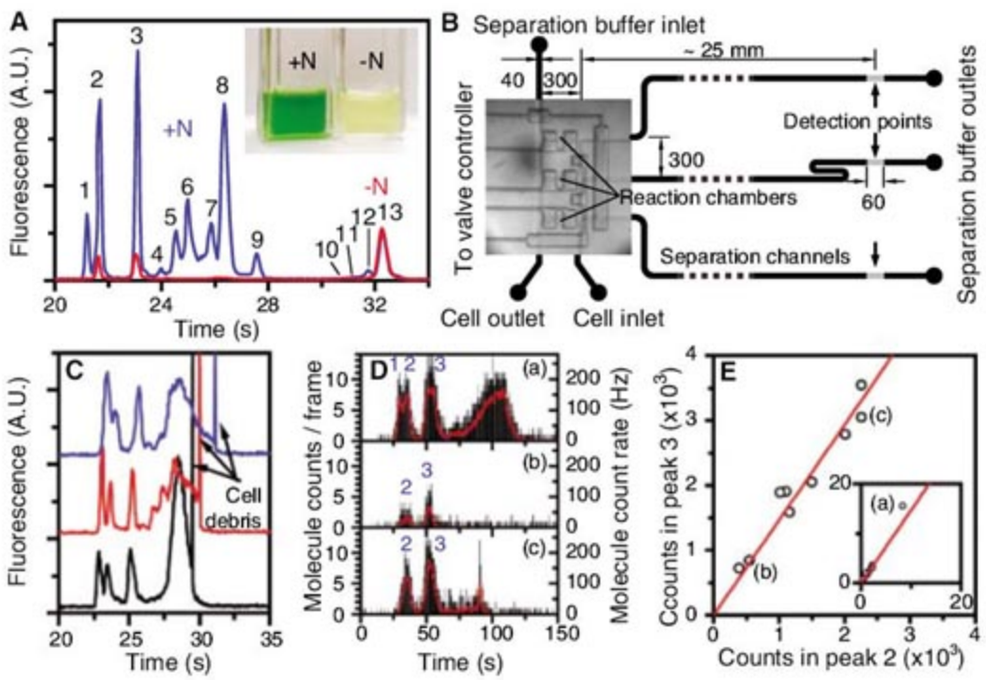


Fig. 3. Single-cell analysis of *Synechococcus* grown in +N and –N culture medium. (**A**) Electropherogram of bulk cell lysates measured in a standard double-T microfluidic chip and normalized to the height of peak 13 (chlorophyll). The inset shows a photograph of the two cell cultures. (**B**) Layout of cyanobacteria analysis chip. Dimensions are in micrometers. (**C**) Electropherograms of phycobiliprotein complexes from three +N cells. The curves are vertically shifted for clarity. (**D**) Molecule-counting results of three –N cells. (**E**) Molecule number distribution of 10 –N cells. The lysing and counting efficiencies are corrected individually. Results from the three cells in (D) are marked by (a), (b), and (c). Red lines show the result of least-square linear fitting. The inset shows cell a, which is excluded from the fitting because its value would dominate the fit.

is the APC-L_{CM} complex and peak 3 is an APC trimer). These observations are consistent with a previously described model for chlorosis (*23*).

Although isolated PC and APC molecules are highly fluorescent, they are difficult to quantify in cells by fluorescence intensities because of the highly efficient energy transfer in the light-harvesting protein complexes, their large spectral overlap, and the fluorescence background from chlorophylls in the photosystems. Our single-cell analysis approach is particularly advantageous because it incorporates chemical separation before fluorescence detection.

We used a chip with three simplified reaction chambers (Fig. 3B) so that up to three cells could be simultaneously lysed. Laser-induced fluorescence detection was used to analyze the +N cells (Fig. 3C) because of their high PBP concentration, whereas molecule counting was used to analyze the –N cells (Fig. 3D). Among the three –N cells shown in Fig. 3D, cell b contains 548 copies of PBP complexes in peak 2 and 839 copies in peak 3, and the detection limit (two times the standard deviation of background counts from noise) is seven molecules. Figure 3E shows the distribution of the molecule numbers of the two PBS core subassemblies (peaks 2 and 3) from 10 –N cells. The cell-to-cell variation in overall PBP counts of –N cells is much larger than that of +N cells. The number of molecules present in these two subassemblies shows a high correlation (r^2 = 0.93) over the entire distribution range. This relation indicates that a constant ratio of these two complexes was maintained during the degradation of the PBS under –N conditions and that, as bleaching of the cells proceeded, the complexes were simultaneously lost. These results suggest coordinated degradation of PBS components within the core of the PBS.

Among the 10 –N cells examined, cell a in Fig. 3D had much brighter fluorescence and much higher molecule counts than the others. Its electropherogram resembled those from +N cells, which indicated an incomplete proteolysis of the PBS. This cell represents those –N cells that have atypically bright fluorescence (~5% of all cells) when observed by fluorescence microscopy. In ensemble experiments in which cell populations are examined, these cells would not be detected because of their low frequencies of appearance. The occurrence of this rare cell could be a consequence of genetic variation within the population, although more work (possibly using *Synechococcus* mutants defective in PBS degradation or using carefully monitored isogenic lines) needs to be done to test this hypothesis.

The method described here should have many applications, making the chemical analysis of single cells by single-molecule counting a new tool for understanding the functioning of cells. Other applications, such as the use of fluorescent proteins that can be genetically fused or the use of dye molecules (such as the biarsenical dye, FlAsH, and its variants) that bind to specific peptide sequences, may further expand the scope of this method to include the monitoring of gene expression and suppression.

References and Notes

1. K. Cottingham, *Anal. Chem.* **76**, 235a (2004).
2. J.-Q. Wu, T. D. Pollard, *Science* **310**, 310 (2005).
3. L. Cai, N. Friedman, X. S. Xie, *Nature* **440**, 358 (2006).
4. J. R. S. Newman *et al.*, *Nature* **441**, 840 (2006).
5. J. Yu, J. Xiao, X. Ren, K. Lao, X. S. Xie, *Science* **311**, 1600 (2006).
6. J. W. Hong, V. Studer, G. Hang, W. F. Anderson, S. R. Quake, *Nat. Biotechnol.* **22**, 435 (2004).
7. H. K. Wu, A. Wheeler, R. N. Zare, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 12809 (2004).
8. S. Hu *et al.*, *Anal. Chem.* **76**, 4044 (2004).
9. A. Castro, F. R. Fairfield, E. B. Shera, *Anal. Chem.* **65**, 849 (1993).
10. D. Chen, N. J. Dovichi, *Anal. Chem.* **68**, 690 (1996).
11. Y. Ma, M. R. Shortreed, H. Li, W. Huang, E. S. Yeung, *Electrophoresis* **22**, 421 (2001).
12. B. Huang, H. Wu, S. Kim, R. N. Zare, *Lab Chip* **5**, 1005 (2005).
13. G. Ocvirk *et al.*, *Electrophoresis* **21**, 107 (2000).
14. B. B. Haab, R. A. Mathies, *Anal. Chem.* **71**, 5137 (1999).
15. W. A. Lyon, S. Nie, *Anal. Chem.* **69**, 3400 (1997).
16. A. Lundqvist, D. T. Chiu, O. Orwar, *Electrophoresis* **24**, 1737 (2003).
17. M. Foquet, J. Korlach, W. R. Zipfel, W. W. Webb, H. G. Craighead, *Anal. Chem.* **76**, 1618 (2004).
18. S. Pennathur, J. G. Santiago, *Anal. Chem.* **77**, 6772 (2005).
19. S. Pennathur, J. G. Santiago, *Anal. Chem.* **77**, 6782 (2005).
20. Materials and methods are available as supporting material on *Science* Online.
21. B. K. Kobilka, *Anal. Biochem.* **231**, 269 (1995).
22. R. J. Whelan *et al.*, *Anal. Chem.* **74**, 4570 (2002).
23. A. R. Grossman, M. R. Schaefer, G. G. Chiang, J. L. Collier, *Microbiol. Rev.* **57**, 725 (1993).
24. A. N. Glazer, D. J. Lundell, G. Yamanaka, R. C. Williams, *Ann. Inst. Pasteur Microbiol.* **B134**, 159 (1983).
25. G. Yamanaka, A. N. Glazer, *Arch. Microbiol.* **124**, 39 (1980).

26. We thank F. S. Thai and X. Yao for the SF9 cell culture, F. Fazeli for the cyanobacteria culture, S. Bailey for purifying the PBS, X. Shi for helping in absorption spectra measurement, and Stanford Nanofabrication Facilities for providing the photolithography equipment. Supported by NSF through grant BES-0508531.

Supporting Online Material

www.sciencemag.org/cgi/content/full/315/5808/81/DC1
Materials and Methods
Figs. S1 to S6
Table S1
References
Movies S1 and S2

17 August 2006; accepted 10 November 2006
10.1126/science.1133992

Mass-Independent Sulfur Isotopic Compositions in Stratospheric Volcanic Eruptions

Mélanie Baroni,[1]* Mark H. Thiemens,[2] Robert J. Delmas,[1] Joël Savarino[1]*

The observed mass-independent sulfur isotopic composition ($\Delta^{33}S$) of volcanic sulfate from the Agung (March 1963) and Pinatubo (June 1991) eruptions recorded in the Antarctic snow provides a mechanism for documenting stratospheric events. The sign of $\Delta^{33}S$ changes over time from an initial positive component to a negative value. $\Delta^{33}S$ is created during photochemical oxidation of sulfur dioxide to sulfuric acid on a monthly time scale, which indicates a fast process. The reproducibility of the results reveals that $\Delta^{33}S$ is a reliable tracer to chemically identify atmospheric processes involved during stratospheric volcanism.

Plinian volcanic eruptions dramatically modify climate for several years by injecting large amounts of dust and gases, such as sulfur dioxide (SO_2), directly into the stratosphere. Therefore, studying the impact of such eruptions can provide insights into atmospheric chemical sensitivity. The sulfuric acid (H_2SO_4) layer that formed from SO_2 oxidation, within about 1 month for moderate eruptions (*1*, *2*), reflects solar radiation and alters the radiative properties of the atmosphere, cooling the global troposphere (*1*). Satellite observations (*3*) or light detection and ranging measurements (*1*) allow recent volcanic events to be monitored, but the existing volcanic observations database is insufficient to represent all past eruptions (*1*), and it remains difficult to include volcanic events as a climatic forcing parameter in efforts to understand past climates and to predict the impact of future eruptions. An important step in understanding the climatic impact of volcanic events of different sizes is to identify the chemical processes linked to stratospheric eruptions. During these events, the chemical composition of the atmosphere is strongly perturbed, especially the sulfur budget. The magnitude and character of the perturbation depend on parameters such as eruption strength, sulfur loading, altitude of the volcanic plume, latitude of the volcano, and time of year.

The sulfate concentration of ice cores has proven to be a useful chemical proxy of past vol-

[1]Laboratoire de Glaciologie et Géophysique de l'Environnement, CNRS/Université Joseph Fourier, 38400 St. Martin d'Hères, France. [2]Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA 92093–0356, USA.

*To whom correspondence should be addressed. E-mail: baroni@lgge.obs.ujf-grenoble.fr (M.B.); jsavarino@lgge.obs.ujf-grenoble.fr (J.S.)

canism (*1*), though such records are restricted in interpretative value. Here we report measurements of the mass-independent sulfur isotopic composition of volcanic sulfate, which provide further insights into atmospheric chemical processes associated with stratospheric volcanism.

Mass-dependent isotopic fractionation processes are governed by relative mass differences between the four sulfur isotopes and are described by $\delta^{33}S \approx 0.515\ \delta^{34}S$ and $\delta^{36}S \approx 1.91\ \delta^{34}S$. The deviation from the two mass-dependent relationships is termed mass-independent fractionation (MIF) and quantified by $\Delta^{33}S$ and $\Delta^{36}S$ (*4*). Sulfur isotopic anomalies have been found in Archean sulfide and sulfate minerals older than 2 billion years (*5*, *6*) but not in younger rocks, suggesting changes in the sulfur oxidation pathways, which are probably linked to the modification in oxygen abundance and signal the appearance of an ozone layer, and in the shielding ultraviolet (UV) light required to generate the sulfur isotopic anomalies at low altitudes (*5*–*7*). A wavelength (λ) dependency of the sulfur MIF has been postulated (*5*), and SO_2 photolysis experiments conducted at different λ have been used to argue this hypothesis (*8*, *9*); but once again, such a view has been recently challenged by new geochemical observations (*10*), although these observations do not formally rule out the shielding UV-light hypothesis. The sulfur isotopic anomaly is generated by UV light that is shorter than $\lambda = 310$ nm, which is only available in the stratosphere in the present atmosphere. Sulfur MIF has been reported in tropospheric aerosols (*11*) and attributed to stratospheric input of sulfate into the troposphere. Stratospheric volcanic eruptions offer the opportunity to study sulfur MIF under natural atmospheric conditions because large quantities of SO_2 are directly injected into the stratosphere and the sulfate is deposited and stably preserved in Antarctic snow. The volcanic signals in ice core records of the stratospheric eruptions of Pinatubo (1991) and an unknown event (1259) exhibit substantial sulfur isotopic anomalies, contrary to the high-tropospheric Cerro Hudson event (1991) (*12*, *13*).

The mechanism responsible for the sulfur isotopic anomaly formation is still under discussion. A recent model study (*14*) predicts a change in the sign of $\Delta^{33}S$ over time, which provides a testable theoretical framework for production of the isotopic anomalies from volcanic events.

Sulfate backgrounds and sulfate deposited on the Antarctic ice sheet by the Agung (8°S, 115°E) and Pinatubo (15°N, 120°E) eruptions in March 1963 and June 1991, respectively, were extracted from snow layers sampled at a high time resolution in a snow pit located at Dome C (75°S, 123°E) and were analyzed for their sulfur isotopic content (*15*). In nonvolcanic conditions, the sulfate deposited in Dome C snow is predominantly biogenic (*16*) and was formed in the troposphere. The backgrounds analyzed in this work confirm that sulfate that was formed out of the stratosphere (i.e., away from photochemical oxidations) was free of any sulfur isotopic anomalies [$\Delta^{33}S \approx 0$ per mil (‰)] (Figs. 1 and 2).

On the other hand, the Agung and Pinatubo volcanic sulfate coming from the stratosphere exhibit large sulfur MIF. In both cases, the maximum sulfur isotopic anomaly occurs before the maximum sulfate concentration (Fig. 1). We observed a stable $\Delta^{33}S$ for about a half year, ~1‰ (corrected for background values; table S2), which was followed by a decreasing anomaly, passing through 0‰ about 1 year after the beginning of the sulfate deposition to reach a bottom value close to –1‰ at the end of deposition. $\delta^{34}S$ values for both Agung and Pinatubo events have an identical range (Fig. 2). The decrease of $\Delta^{33}S$ from positive to negative values is accompanied by a significant decrease of $\delta^{34}S$ over time.

During the beginning of the Pinatubo eruption, the $\Delta^{33}S$ and $\delta^{34}S$ values (Fig. 2) are low with respect to the Agung profile. These values are the result of a mass-dependent process caused by the input from the Cerro Hudson (46°S, 73°W) eruption. This volcano erupted in August 1991 and injected 2 Tg of SO_2 up to the upper troposphere–lower stratosphere region (*13*) (that is, below an altitude of 14 km). At this altitude, the lack of UV radiation shorter than $\lambda = 310$ nm does not allow SO_2 photolysis, which is responsible for the generation of the sulfur isotopic anomaly (*8*). Measurements of the Cerro Hudson sulfate recorded in snow at the South Pole site (90°S, 102°W), where snow accumulation (~20 cm $year^{-1}$) permits analysis of the two volcanic events independently, have confirmed the sulfur mass-dependent nature of the Cerro Hudson sulfate, which exhibits a near-zero $\Delta^{33}S$ and a $\delta^{34}S$ of +5.60‰ (*12*). Because of its geographic location, the Cerro Hudson plume reached Antarctica before that of Pinatubo (*17*, *18*). The two volcanic signals are not separated in Dome C snow because of the low snow-accumulation rate (~10 cm $year^{-1}$) at this site. Consequently, the $\Delta^{33}S$ and $\delta^{34}S$ values (+0.38 and +10.37‰, respectively) correspond to the mass-independent Pinatubo signal that was diluted by the mass-dependent Cerro Hudson contribution.

The mean sulfur isotopic composition of sulfate recorded in the snow during volcanic deposition was estimated with the use of a weighted average. The mean values of $\delta^{34}S$ and $\Delta^{33}S$ are +9.57 and +0.70‰ for Agung and +9.29 and +0.29‰ for Pinatubo. For the Pinatubo record, the point at 135 cm is not considered in the weighted average because of the mass-dependent nature of the Cerro Hudson contribution. These calculations allow for the comparison with measurements obtained at the South Pole (*12*, *19*). The Agung $\delta^{34}S$ mean signal measured at the South Pole (*19*) is estimated to be 2.7 ± 1.1‰, a low value as compared to our calculation. This difference is difficult to explain, but the Pinatubo record can provide some elements for evaluation. The Pinatubo $\delta^{34}S$ and $\Delta^{33}S$ values at Dome C are lower than the ones measured at the South Pole (+10.89 and +0.67‰, respectively) (*12*). The South Pole measurements were made only during the first year of sulfate deposition when the isotopic anomaly is positive and the sulfur isotopic composition is enriched in heavy isotopes as compared to the end of deposition. Consequently, fractional analysis of the South Pole Pinatubo record can give higher values as compared to what they would be if the total sulfate deposition was analyzed. Restricting our calculation of the weighted mean to the same chronology, we obtain 13.12 and 0.69‰ for the $\delta^{34}S$ and $\Delta^{33}S$ values, respectively, which is consistent with the South Pole measurements.

Lithospheric volcanic sulfur does not have a sulfur isotopic anomaly; thus, the isotopic budget requires a mass-weighted positive $\Delta^{33}S$ created in the stratosphere to be balanced by a mass-weighted negative $\Delta^{33}S$. We observe a change in the sign of $\Delta^{33}S$, but the budget for both eruptions is significantly different than 0, probably because a portion of the sulfate is lost before reaching Antarctica. The signal loss preferentially affects

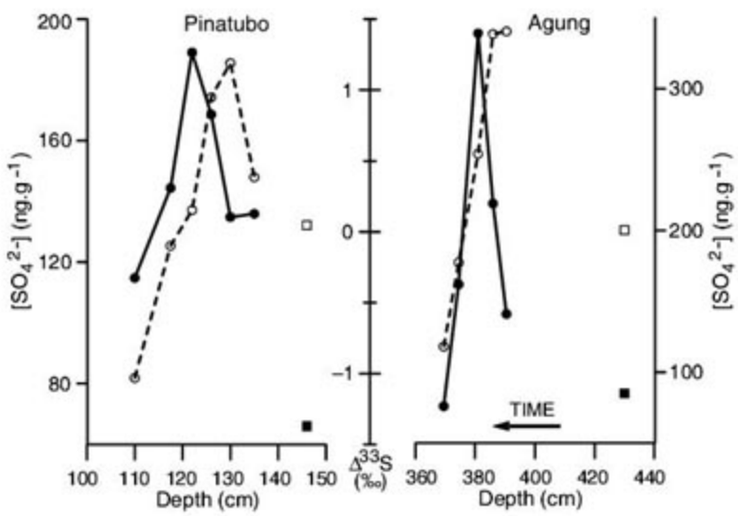


Fig. 1. Plot of sulfate concentrations (solid symbols) and $\Delta^{33}S$ (open symbols) as a function of depth and, by extension, as a function of time for both the Agung and Pinatubo volcanic eruptions (circles). Squares represent isotopic and concentration background data; the background values reported in the present work are consistent with those of previous studies (*12*, *26*). These eruptions are the source of the two major volcanic sulfate deposits of the past 50 years, which are easily accessible from Antarctic snow. The maximum $\Delta^{33}S$ value precedes the maximum sulfate concentration. The $\Delta^{33}S$ value changes in sign over time, starting with a positive phase.

the negative $\Delta^{33}S$ component because the snow budget has a positive average value.

The isotopic evolution of the volcanic sulfate is substantial. For the Agung record, the $\delta^{34}S$ fractionation attains 17.5‰ between the beginning and end of the eruption. Such fractionation was observed in stratospheric aerosols at 40°S after the Agung eruption (*20*). The aerosol measurements were reported from October 1962 to August 1971, but we focused on the period covering the Agung eruption (May 1963 to June 1965). In May 1963, the Agung eruption was complete, and nearly all SO_2 emitted had already been oxidized to H_2SO_4. The slopes of $\delta^{34}S$ plotted as a function of time (Fig. 3) show that the isotopic fractionation was more important at an altitude of 18.3 km than at 15.2 km, which suggests a gradient in isotopic composition with altitude. The similarity between the slopes representing stratospheric aerosols at 15.2-km altitude and sulfate deposited in the Dome C snow makes us think that transport from the tropopause to the ground and depositional effects do not change the $\delta^{34}S$ notably. The sulfate was measured at different latitudes (40°S for stratospheric aerosols and 75°S for snow samples), which suggests that no substantial change in the sulfur isotopic composition occurs during the transport into the stratosphere. The fractionation that is observed as a function of time was attributed to an inverse kinetic-isotope effect, where $^{34}SO_2$ + OH reacts faster than $^{32}SO_2$ + OH, producing enriched ^{34}S sulfate first, followed by gradually depleted aerosols over time (*20*, *21*). This process is mass dependent and alters $\delta^{34}S$ but not $\Delta^{33}S$ in the range of our observed values, even if individual processes may have a slightly different mass-dependent exponent. Because $\delta^{34}S$ does not appear to undergo fractionation during transport, we may extrapolate this statement to the isotopic sulfur anomaly because $\Delta^{33}S$ and $\delta^{34}S$ are linearly correlated (Fig. 2). Consequently, H_2SO_4 does not appear to undergo any process that alters its sulfur isotopic composition during transport and subsequently its mass-independent composition. The most plausible reaction scenario is that the sulfur isotopic anomaly is generated from SO_2 before its oxidation to H_2SO_4.

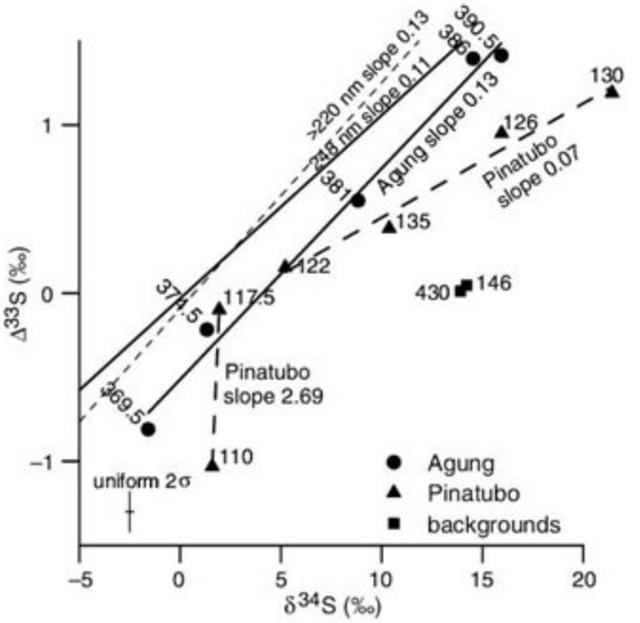


Fig. 2. Plot of $\Delta^{33}S$ versus $\delta^{34}S$ for both the Agung and the Pinatubo eruptions and comparison with SO_2 photolysis experiments conducted at different λ (*8*, *9*). The numbers associated with the points are the depths in snow. A uniform error bar (2σ) is indicated on the graph. The point at a depth of 135 cm can be disregarded because of the bias created by the Cerro Hudson eruption. The change in sign of $\Delta^{33}S$ over time is accompanied by a decrease of the $\delta^{34}S$ values. The $\Delta^{33}S$ versus $\delta^{34}S$ results from Pinatubo differ for those from Agung. The Pinatubo points at 117.5 and 110 cm are of special interest because they have negative $\Delta^{33}S$ and $\Delta^{36}S$ as compared to all other samples measured in this work (table S2) and in most of the Archean rocks (*5*, *6*) or other volcanic samples (*12*). This feature was only observed during SO_2 photolysis carried out at 184.9 nm (*8*). This wavelength is not available at the altitude attained by the two volcanic plumes. However, the Pinatubo plume reached higher altitudes than that of Agung; thus the SO_2 could have photolysed at shorter wavelengths, perhaps accounting for the discrepancy between the two volcanic events, but such a hypothesis needs to be tested by new SO_2 photolysis experiments.

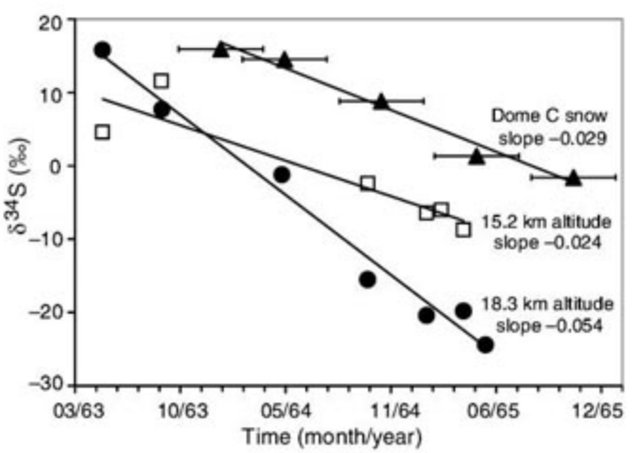


Fig. 3. Plot of $\delta^{34}S$ as a function of time after the Agung eruption. A comparison is shown between the stratospheric aerosols measured at altitudes of 15.2 km (open squares) and from 18.3 to 19.2 km (solid circles) at 40°S (*20*) and the sulfate recorded in the Dome C (75°S) snow (solid triangles) (this work). Error bars indicate an estimated age uncertainty based on snow accumulation uncertainty. The age used for Agung signal recorded in the Dome C snow is based on a snow accumulation of 10 cm $year^{-1}$. The similarity between the slopes obtained from snow and aerosol samples taken at different altitudes and latitudes suggests that the transport does not significantly affect the sulfate isotopic composition.

Sulfur dioxide photodissociation below 220 nm [(SO_2 + $h\nu$ = SO + O), where $h\nu$ is photonic energy] accounts for the sulfur MIF in the Archean anoxic atmosphere (*7*, *14*) but not in today's atmosphere because sulfur monoxide (SO) rapidly reacts with O_2 to re-produce SO_2 (*12*, *14*). Sulfur dioxide photooxidation ($SO_2^* + SO_2 \rightarrow SO_3 + SO$) has been proposed as a mechanism to generate the observed sulfur isotopic anomaly (*12*). Excited SO_2 can react with ground-state SO_2 when its number densities are anomalously high, such as in the first weeks subsequent to a massive volcanic eruption. This reaction may occur at lower altitudes of the stratosphere as compared to SO_2 photodissociation if the SO_2 plume remains optically thin in the UV window ($260 < \lambda < 340$ nm). Laboratory experiments have also confirmed the ability of photolysis at these wavelengths to produce sulfur MIF species (*8*, *9*). The creation of two reservoirs composed of SO_2 and sulfur trioxide (SO_3), which can carry sulfur MIF with the opposite sign, is the requisite condition to generate a $\Delta^{33}S \neq 0‰$ in the volcanic sulfate. According to laboratory experiments conducted at $\lambda > 220$ nm (*8*, *9*), residual SO_2 has a $\Delta^{33}S < 0‰$, and produced sulfate (probably via SO_3) has a $\Delta^{33}S > 0‰$. Because, in the atmosphere, the reaction to produce H_2SO_4 between SO_3 and H_2O is faster than that between SO_2 and OH, it is anticipated that sulfate will initially carry a positive $\Delta^{33}S$ and will later become negative as a result of material-balance considerations. Such a view is consistent with the field observations. Recently, SO_2 photooxidation was questioned, and SO_3 photolysis ($SO_3 \rightarrow SO_2$ + O) was invoked to explain sulfur MIF (*14*). As in the case of SO_2 photooxidation, SO_3 photolysis may occur at low altitudes in the stratosphere ($195 < \lambda < 300$ nm) and generate two reservoirs. Based on this hypothesis, it was suggested that the Pinatubo event should produce a change in the sign of the isotopic anomaly of sulfate but with the negative component reaching the ground first (*14*), therefore forcing the near-zero negative $\Delta^{33}S$ of Cerro Hudson to be the beginning of the Pinatubo event, in contradiction with the present isotopic observations and concentration measurements (*17*). SO_3 photolysis dominates at altitudes above 37 km (*22*). Below this altitude, the reaction between SO_3 and H_2O is faster, and SO_3 photolysis is a minor process in the production of a substantial amount of sulfate. SO_2 photooxidation remains the leading candidate to account for sulfur MIF of volcanic sulfate. SO_2 photooxidation as the source of sulfur MIF is in accordance with SO_2 photolysis

experiments (*8*, *9*). The Agung data are also consistent with laboratory experiments because the $\Delta^{33}S$ versus $\delta^{34}S$ Agung slope (Fig. 2) is the same as that of the Xe lamp experiment obtained for $\lambda > 220$ nm and very close to that of the KrF laser experiments conducted at 248 nm (*8*, *9*).

The sulfur isotopic anomalies in volcanic samples are much smaller than those observed in Archean rocks older than 2.45 billion years (*5*, *6*, *23*, *24*). In today's atmosphere, OH radicals remain the main sink of SO_2 emitted after a volcanic eruption, and the $SO_2^* + SO_2$ reaction is a minor reaction when compared to the $SO_2 + OH$ reaction. The sulfur MIF measured in volcanic sulfate recorded in snow is a diluted signal and may actually reach the extreme values recorded in Archean rocks. To estimate the upper limit of the sulfur isotopic anomaly generated by the photooxidation process, researchers should compare the kinetics of the $SO_2 + OH$ and $SO_2^* + SO_2$ reactions. Unfortunately, the rate of $SO_2^* + SO_2$ is controversial (*25*) and is needed for such quantification.

Sulfur mass-independent composition of volcanic sulfate is a time-dependent process, first displaying a positive $\Delta^{33}S$ followed by a negative $\Delta^{33}S$ at the end of the volcanic plume depositional process. This process occurs on a monthly time scale before SO_2 is fully oxidized in H_2SO_4, indicating a rapid process. The nonzero average $\Delta^{33}S$ observed for the full duration of the event requires two conditions: First, the process creates two reservoirs of MIF with opposing signs; second, these two reservoirs must be physically separated in space and time in addition to having a difference in depositional rates. The only way to explain the oscillation of the $\Delta^{33}S$ sign is to consider the fundamental role of aerosols and sedimentation in preserving the isotopic signal. Microphysical processes must be taken into account in models to reproduce sulfur MIF of stratospheric volcanic sulfate. When the relationship between aerosols and sulfur MIF is established, volcanic plume transport may be understood, allowing a precise glaciological record of the climatic impact of stratospheric eruptions.

References and Notes

1. G. Zielinski, *Quat. Sci. Rev.* **19**, 417 (2000), and references therein.
2. S. Bekki, J. A. Pyle, *J. Geophys. Res.* **99**, 18861 (1994).
3. G. J. S. Bluth, C. C. Schnetzler, A. J. Krueger, L. S. Walter, *Nature* **366**, 327 (1993).
4. The deviation from the mass-dependent relationships is calculated by the following equations: $\Delta^{33}S = \delta^{33}S - 1000*[(1 + \delta^{34}S/1000)^{0.515} - 1]$ and $\Delta^{36}S = \delta^{36}S - 1000*[(1 + \delta^{34}S/1000)^{1.91} - 1]$. Considering the small size of the samples, our analytical accuracy, with a 2σ uncertainty, is equal to 0.12‰ for $\Delta^{33}S$ and varies from 0.64 to 1.63‰ for $\Delta^{36}S$. Only $|\Delta^{33}S| > 0.12$‰ and $|\Delta^{36}S| > 0.64$‰ are considered as diagnostic of MIF in the present study. Uncertainties (2σ) are 0.07, 0.19, and 0.53 to 1.59‰ for $\delta^{33}S$, $\delta^{34}S$, and $\delta^{36}S$, respectively.
5. J. Farquhar, H. Bao, M. H. Thiemens, *Science* **289**, 756 (2000).
6. J. Farquhar, B. A. Wing, *Earth Planet. Sci. Lett.* **213**, 1 (2003).
7. A. A. Pavlov, J. F. Kasting, *Astrobiology* **2**, 27 (2002).
8. J. Farquhar, J. Savarino, S. Airieau, M. H. Thiemens, *J. Geophys. Res.* **106**, 32829 (2001).
9. J. Farquhar, J. Savarino, T. L. Jackson, M. H. Thiemens, *Nature* **404**, 50 (2000).
10. H. Ohmoto, Y. Watanabe, H. Ikemi, S. R. Poulson, B. E. Taylor, *Nature* **442**, 908 (2006).
11. A. Romero, M. H. Thiemens, *J. Geophys. Res.* **108**, 10.1029/2003JD003660 (2003).
12. J. Savarino, A. Romero, J. Cole-Dai, S. Bekki, M. H. Thiemens, *Geophys. Res. Lett.* **30**, 10.1029/2003GL018134 (2003).
13. S. D. Doiron, Z. Scott, G. J. S. Bluth, C. C. Schnetzler, A. J. Krueger, *EOS Trans. Am. Geophys. Union* **72**, 489 (1991).
14. A. A. Pavlov, M. J. Mills, O. B. Toon, *Geophys. Res. Lett.* **32**, 10.1029/2005GL022784 (2005).
15. Materials and methods are available as supporting material on *Science* Online.
16. M. R. Legrand, R. J. Delmas, *Atmos. Environ.* **18**, 1867 (1984).
17. J. Cole-Dai, E. Mosley-Thompson, *Ann. Glaciol.* **29**, 99 (1999).
18. M. Legrand, D. Wagenbach, *J. Geophys. Res.* **104**, 1581 (1999).
19. N. Patris, R. J. Delmas, J. Jouzel, *J. Geophys. Res.* **105**, 7071 (2000).
20. A. W. Castleman Jr., H. R. Munkelwitz, B. Manowitz, *Tellus* **26**, 222 (1974).
21. F.-Y. Leung, A. J. Colussi, M. R. Hoffmann, *J. Phys. Chem. A* **105**, 8073 (2001).
22. J. B. Burkholder, S. McKeen, *Geophys. Res. Lett.* **24**, 3201 (1997).
23. S. Ono *et al.*, *Earth Planet. Sci. Lett.* **213**, 15 (2003).
24. S. J. Mojzsis, C. D. Coath, J. P. Greenwood, K. D. McKeegan, T. M. Harrison, *Geochim. Cosmochim. Acta* **67**, 1635 (2003).
25. K. Chung, J. G. Calvert, J. W. Bottenheim, *Int. J. Chem. Kinet.* **7**, 161 (1975).
26. E. Castellano *et al.*, *J. Geophys. Res.* **110**, 10.1029/2004JD005259 (2005).

27. We acknowledge useful discussions with S. Bekki and thank the Conseil Régional Rhônes-Alpes for partially supporting travel expenses for M.B. J.S. acknowledges the Balzan Foundation for financial support; C. Lorius and the Institut National des Sciences de l'Univers (INSU) for mass spectrometry acquisition; the French Polar Institute and M. Legrand (Institut Polaire Francais Paul Emile Victor program DC17) for logistical support in Antarctica; the CNRS, under its Programme International de Coopération Scientifique; and the INSU Programme National de Chimie Atmosphérique. The NSF Office of Polar Programs provided financial support for M.H.T. We also thank J. McCabe and U. Morgenstern for helping J.S. to dig the snow pit.

Supporting Online Material

www.sciencemag.org/cgi/content/full/315/5808/84/DC1
Materials and Methods
SOM Text
Fig. S1
Tables S1 and S2
References

26 June 2006; accepted 15 November 2006
10.1126/science.1131754

CO_2-Forced Climate and Vegetation Instability During Late Paleozoic Deglaciation

Isabel P. Montañez,[1]* Neil J. Tabor,[2] Deb Niemeier,[3] William A. DiMichele,[4] Tracy D. Frank,[5] Christopher R. Fielding,[5] John L. Isbell,[6] Lauren P. Birgenheier,[5] Michael C. Rygel[5]†

The late Paleozoic deglaciation is the vegetated Earth's only recorded icehouse-to-greenhouse transition, yet the climate dynamics remain enigmatic. By using the stable isotopic compositions of soil-formed minerals, fossil-plant matter, and shallow-water brachiopods, we estimated atmospheric partial pressure of carbon dioxide (pCO_2) and tropical marine surface temperatures during this climate transition. Comparison to southern Gondwanan glacial records documents covariance between inferred shifts in pCO_2, temperature, and ice volume consistent with greenhouse gas forcing of climate. Major restructuring of paleotropical flora in western Euramerica occurred in step with climate and pCO_2 shifts, illustrating the biotic impact associated with past CO_2-forced turnover to a permanent ice-free world.

A decade of studying Pleistocene ice cores has unequivocally documented a strong coupling of atmospheric partial pressure of CO_2 (pCO_2) and surface temperatures with changing global ice volume (*1*, *2*). Although the precise mechanistic link between atmospheric greenhouse gases and climate is debated, there remains little doubt that high concentrations of atmospheric CO_2 have strongly amplified Earth's past climates. Anthropogenic CO_2 emissions have increased atmospheric CO_2 to concentrations higher than at any time in at least the past 650,000 years and could increase it to more than 2000 parts per million by volume (ppmv) as accessible fossil fuel reservoirs are exhausted (*3*). The last time such concentrations were seen on Earth was at the onset of our modern icehouse [~40 to 34 million years ago (Ma)], a transition from ice-free to glacial conditions characterized by repeated C cycle perturba-

[1]Department of Geology, University of California, Davis, CA 95616, USA. [2]Department of Geological Sciences, Southern Methodist University, Dallas, TX 75275, USA. [3]Department of Civil and Environmental Engineering, University of California, Davis, CA 95616, USA. [4]Department of Paleobiology, Smithsonian Museum of Natural History, Washington, DC 20560, USA. [5]Department of Geosciences, 214 Bessey Hall, University of Nebraska, Lincoln, NE 68588, USA. [6]Department of Geosciences, University of Wisconsin, Post Office Box 413, Milwaukee, WI 53201, USA.

*To whom correspondence should be addressed. E-mail: montanez@geology.ucdavis.edu
†Present address: Department of Geology, State University of New York, College at Potsdam, Potsdam, NY 13676, USA.

tion, large magnitude changes in atmospheric pCO_2, and major ephemeral warmings (*4*, *5*). As our climate system departs from the well-studied Pleistocene glacial-interglacial cycles, a deep-time perspective of pCO_2-climate-glaciation linkages is essential for a fuller understanding of what may be the Earth's most epic deglaciation.

We present here the results of a multipronged investigation that provides evidence for significantly changing atmospheric CO_2 concentrations and surface temperatures during a 40-million-year period of the late Paleozoic (~305 to 265 Ma), which encompasses the deterioration of the most widespread and long-lived icehouse of the last half-billion years (*6*). This global warming event accompanied a permanent transition to an ice-free world, a condition that arguably lasted until the current glacial state. These results, when integrated with a newly emerging glaciation history for southern Gondwana (*7*–*11*), indicate strong linkages between pCO_2, climate, and ice-mass dynamics during the final stages of the Late Paleozoic Ice Age (end of LPIA). Integration of these climate proxy records with our newly developed tropical paleobotanical records shows repeated climate-driven ecosystem restructuring in western paleoequatorial Euramerica.

The CO_2 contents of ancient atmospheres can be estimated from the carbon stable isotope values ($\delta^{13}C$) of ancient soil-formed carbonates and goethites with an uncertainty of $\leq\pm500$ ppmv (*12*, *13*). These minerals are the proxy of choice when pCO_2 is high ($\geq$1000 ppmv), whereas the method's sensitivity decreases at lower pCO_2 (<800 ppmv) (*14*, *15*). The precision of pCO_2 estimates reflects the variable assumptions used for each pCO_2 calculation (*16*), which can be further refined if the $\delta^{13}C$ of coexisting organic matter is available and if quantitative estimates of paleosoil-respired CO_2 content and paleotemperatures can be inferred from modern analogs or independently derived geochemical proxies (*15*).

To reconstruct atmospheric CO_2 during the end of the LPIA, we measured the $\delta^{13}C$ values of soil-formed calcites ($\delta^{13}C_{carb}$) collected from mature, well-drained profiles from the Eastern Shelf of the Midland Basin; the Pedregosa, Anadarko, and Paradox Basins; and the Grand Canyon Embayment of western paleoequatorial Euramerica (fig. S1 and table S1) (*17*). We consider measured paleosol $\delta^{13}C_{carb}$ values to be a robust proxy of soil-water CO_2 during formation, given the lack of evidence for mineral recrystallization and overgrowth and their overall shallow and low-temperature burial histories (*18*). Furthermore, we consider the $\delta^{13}C$ of well-preserved fossil plant matter ($\delta^{13}C_{org}$) to be a faithful proxy of the C isotope composition of soil-respired CO_2 and, in turn, of atmospheric CO_2 (*19*, *20*). Compression and permineralized fossil plants, cuticles, coal, and charcoal were collected from mudstone deposits of abandoned

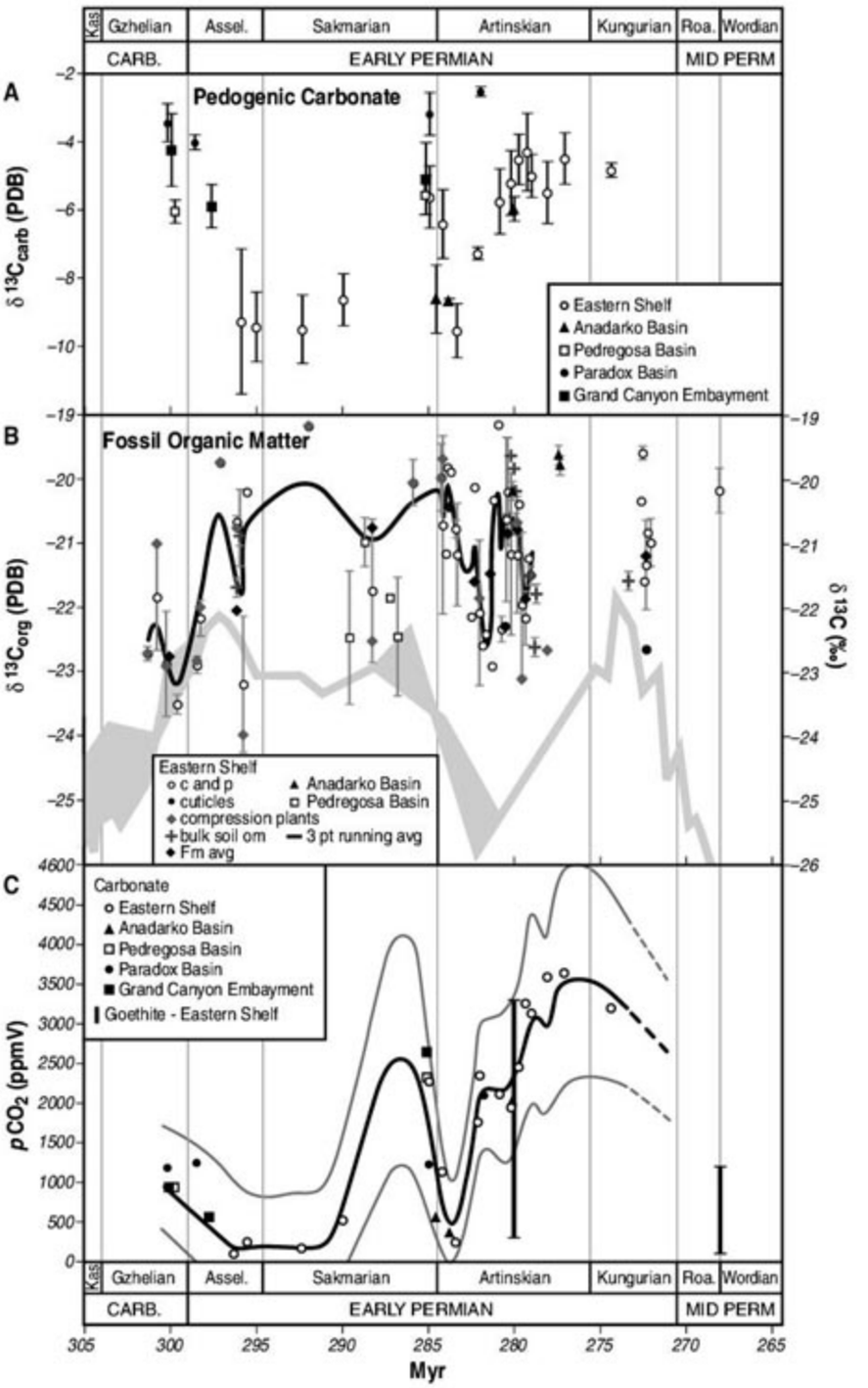


Fig. 1. Temporal distribution of carbonate (**A**) and fossil plant (**B**) $\delta^{13}C$ values used to construct best estimate of Permo-Carboniferous atmospheric pCO_2 (**C**). Individual points in (A) and (B) are the average of analyses from suites of contemporaneous paleosols (from 5 to 18) and associated plant localities (from 3 to 21); "c and p" encompasses all compression and permineralized plant matter, coals, and charcoals. Vertical bars are ±2 SE around the mean. PDB, Pee Dee belemnite. (B) Solid curve is three-point weighted running average through samples from the Eastern Shelf, Midland Basin. Gray band is $\delta^{13}C_{org}$ of Permo-Carboniferous coals from three correlated successions in North China Platform (*22*). Overlapping $\delta^{13}C_{org}$ trends but different $\delta^{13}C_{org}$ values are interpreted to reflect overall wetter conditions for the North China Platform relative to western paleoequatorial Euramerica in the Permian. Data and pCO_2 presented on an age model (*51*) developed for the terrestrial composite section by linearly interpolating between known biostratigraphic boundaries. (C) Best estimate of paleo-pCO_2 (black curve) from Monte Carlo simulation of chronostratigraphically well-constrained sample populations; uncertainty in pCO_2 estimates (gray curves) reflects variability in $\delta^{13}C_{carb}$ and $\delta^{13}C_{org}$, interpreted to record inter- and intrabasinal variations in soil conditions, vegetation, and climate. Vertical bars are published goethite-based CO_2 estimates from the same set of paleosols (*25*).

fluvial channels and floodplains, which are stratigraphically intercalated (on a sub-10-m resolution) with carbonate-bearing paleosols (table S2). The use of measured $\delta^{13}C_{org}$ rather than penecontemporaneous marine carbonates as a proxy of atmospheric $\delta^{13}C$ reflects a growing appreciation of local-scale C cycling effects on the $\delta^{13}C$ values of epicontinental marine carbonates (*21*). The terrestrial $\delta^{13}C_{carb}$ and $\delta^{13}C_{org}$ time series have an average sampling interval of <1 million years (My) and define long-term trends that exhibit systematic variability (Fig. 1, A and B). That the long-term $\delta^{13}C_{org}$ trend records first-order variations in atmospheric $\delta^{13}C$ is supported by its similarity to time-equivalent $\delta^{13}C_{org}$ records of Permo-Carboniferous coals from the North China Platform (*22*) and by a narrow range, throughout the study area, in the ratio of intracellular, *pi*, and atmospheric, *pa*, partial pressures of CO_2 in paleoflora [0.46 to 0.57 ± 0.3 (2 SE)], which were estimated by using measured $\delta^{13}C_{org}$ values of fossil plants and $\delta^{13}C_{carb}$ values of contemporaneous marine brachiopods (*17*). These factors indicate that changes in geomorphic or environmental conditions in the study area were secondary to atmospheric $\delta^{13}C$ in influencing measured fossil-plant $\delta^{13}C_{org}$ values.

Ranges of paleosoil-respired CO_2 content were inferred from the morphologies of suites of contemporaneous paleosols (*23*) by comparison with modern analogs, addressing a major source of uncertainty in previous applications of the CO_2 paleobarometer (table S3) (*14*, *15*). Paleosol temperatures were inferred from the oxygen and hydrogen isotopic compositions of pedogenic phyllosilicates and Fe oxides obtained from the same set of paleosols (*18*, *24*). The best estimate of paleoatmospheric pCO_2 was defined by using Monte Carlo simulation involving 1000 randomly drawn samples for each variable for each time-location combination (*17*). Monte Carlo simulation uses random sampling techniques to stochastically solve physical process problems, in this case quantitatively estimating paleo-pCO_2 and the associated uncertainty by integrating across all of the inferred and measured input variables.

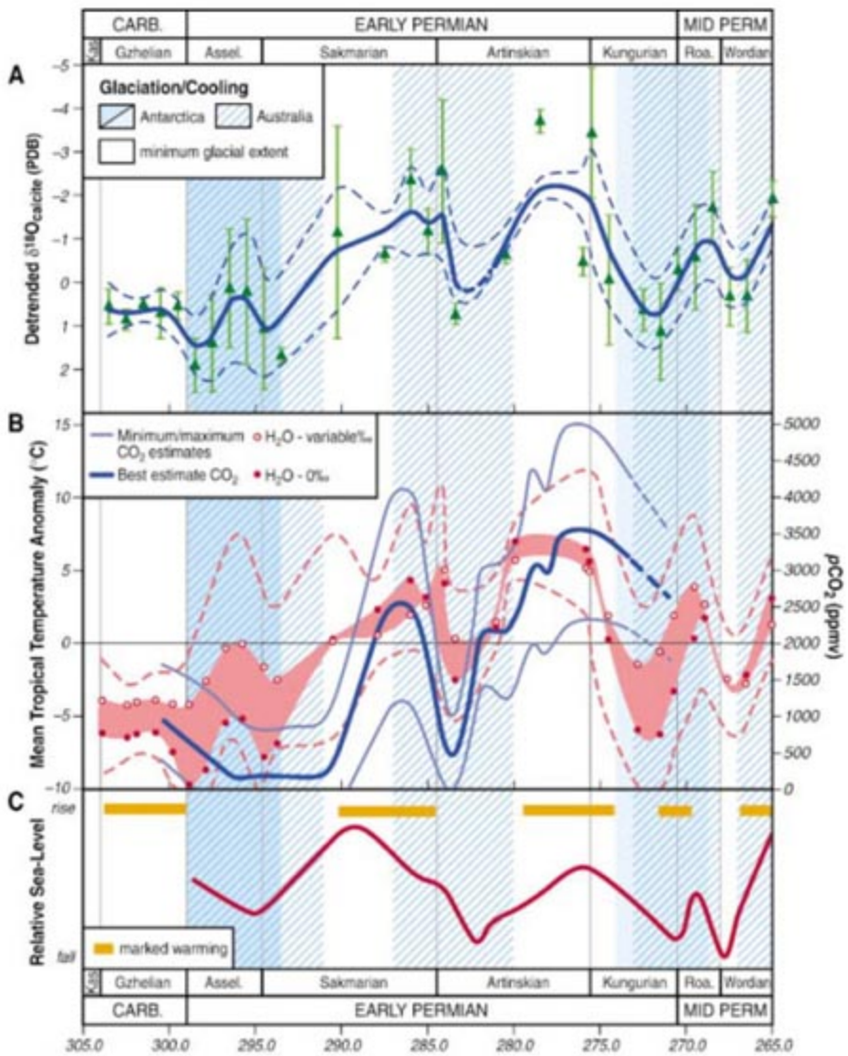


Fig. 2. Relationship among Permo-Carboniferous pCO_2, climate, and cryosphere. Temporal distribution of glacial maxima and/or cool periods based on stratigraphic distribution of diamictites, rhythmites, and dropstone and keel turbate structures in Antarctica and Australian glacigenic deposits (*10*, *11*). (**A**) Three-point weighted running average (blue curve) and ±2 SE (dashed curves) of detrended $\delta^{18}O_{brachiopod}$ values binned into 1- to 3-My windows (green triangles). Error bars indicate ±2 SE around the mean $\delta^{18}O_{brachiopod}$ values. (**B**) Inferred paleotropical SSTs (red interval) (*40*) are reported as temperature anomalies given the potential effects of local and regional environmental and diagenetic influences on brachiopod $\delta^{18}O$. Paleo-SST anomalies (relative to 17.5° C) were calculated from a three-point weighted running average (± 2 SE) through $\delta^{18}O$-based paleotemperature estimates (table S5). Blue curves are best estimate (heavy) and uncertainty (light) of paleo-pCO_2. (**C**) Relative sea-level curve compiled from (*8*, *53*); distribution of warm intervals, from (*7*–*9*) and (*34*).

Modeled CO_2 concentrations (Fig. 1C and table S4) define a long-term rise from an average of present atmospheric levels (PAL = 280 ppmv) in the earliest Permian to values of up to 3500 ppmv by the late Early Permian. A substantial decline in pCO_2 into the early Middle Permian is corroborated by independently derived goethite-based estimates of Permian pCO_2 (*25*). A short-lived (~2 My) drop in pCO_2 to near PAL, defined by contemporaneous paleosols, punctuates the Early Permian rise. Modeled pCO_2 suggests that PAL values were limited to the earliest Permian after latest Carboniferous levels of up to 1000 ppmv, in accord with pCO_2 inferred from marine carbonate $\delta^{13}C$ (*26*) and with southern Gondwanan sedimentologic and geochemical evidence for latest Carboniferous warming (*9*, *27*). Our record refines the structure of well-established pCO_2 reconstructions, which indicate sustained PAL values throughout much of the Permo-Carboniferous (*15*, *28*, *29*). The higher-frequency oscillations revealed by this study would be below the temporal resolution (5 to 20 My time-averaging) of those long-term CO_2 records.

In order to evaluate the nature of the CO_2-climate relationship, we developed a time-equivalent record of paleotropical sea-surface temperatures (SSTs) by using $\delta^{18}O$ values from a global compilation of well-preserved latest Carboniferous through Middle Permian tropical shallow-water brachiopods (table S5) (*30*); brachiopods have diagenetically resistant, low-Mg calcitic shells that incorporate oxygen isotopes in equilibrium with seawater (*31*). The residual brachiopod $\delta^{18}O$ record (Fig. 2A) displays clear isotopic fluctuations, with intervals of maximum values corresponding to Permian glacial maxima or marked coolings in Antarctica and/or Australia (*10*, *11*) and, to the degree afforded by geochronologic dates, with the younger periods of inferred glacial maxima in the Karoo Basin (*8*, *32*), southern Argentina (*9*), and Tasmania (*33*). Intervals of minimum $\delta^{18}O_{carb}$ values correspond with independently inferred periods of marked warming and sea-level rise (*7*–*9*, *34*) (Fig. 2C).

Inferring secular paleotemperatures from $\delta^{18}O_{carb}$ requires careful consideration of the compound effects on values of continental ice

volume, local hydrography, and SST, as well as any vital effects and postdepositional alteration (*31*, *35*). The eustatic component in the Permo-Carboniferous brachiopod $\delta^{18}O$ record due to ice volume variability likely accounts for far less than 2 per mil (‰) of the observed $\delta^{18}O$ variation given reconstructed amplitudes (10 to <100 m) of Permo-Carboniferous glacio-eustasy (*10*) and an O isotope composition of seawater ($\delta^{18}O_{sw}$)–sea level relationship of 0.1‰ per 10 m of sea level change (*36*). The residual secular $\delta^{18}O_{carb}$ signal is interpreted to record changes in temperature, salinity, and pH. Local hydrographic variations in tropical epicontinental seas would have dampened the magnitude of $\delta^{18}O_{carb}$ shifts, given hypothesized heightened freshwater discharge to continental shelves (decreased salinity and lowered $\delta^{18}O_{sw}$) during late Paleozoic periods of maximum glaciation, and increased evaporation (increased salinity and $\delta^{18}O_{sw}$) during drier, highly seasonal glacial minima (*36*). Moreover, paleo-SSTs under elevated pCO_2 may be underestimated by up to 2°C, given that lowered seawater pH would have shifted $\delta^{18}O_{carb}$ to less negative values (*38*, *39*).

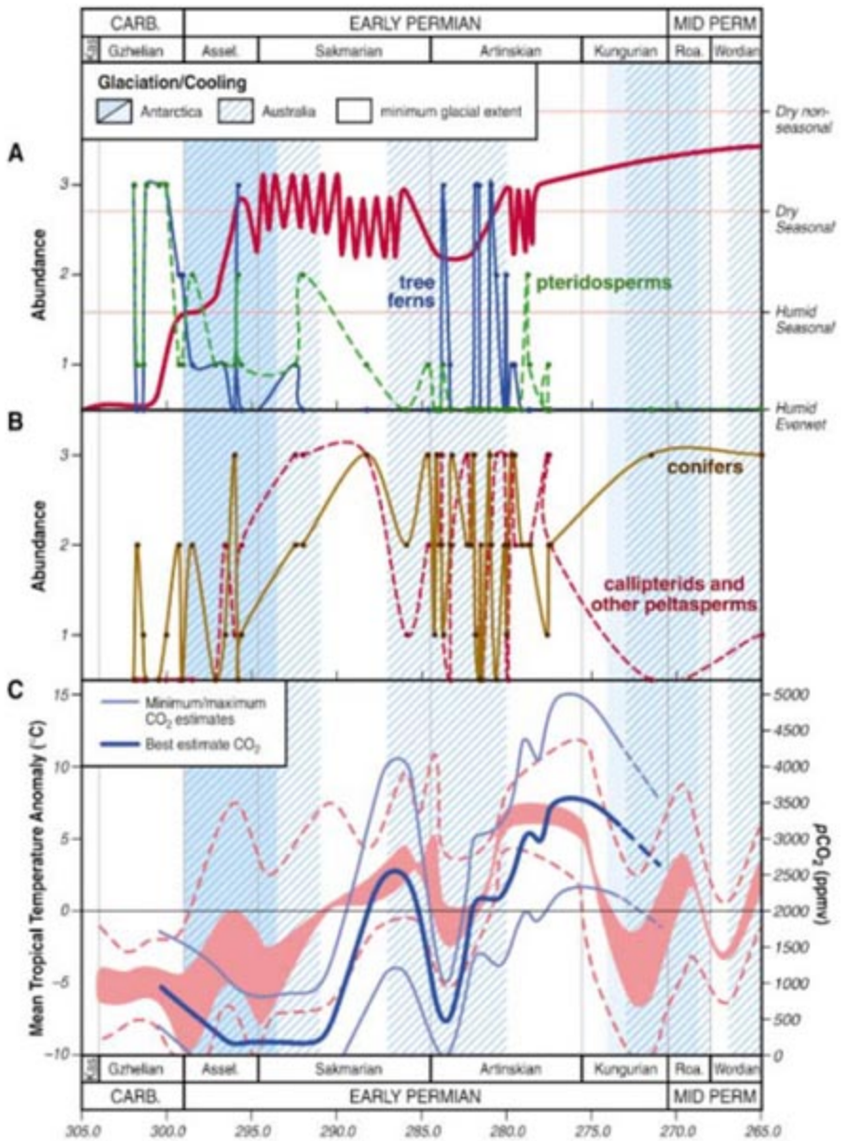


Fig. 3. Patterns of abundance change in major flora of study area (**A** and **B**) and comparison to independently derived Permo-Carboniferous climate and pCO_2 (**C**). Plants from 49 sampling localities on the Eastern Shelf, Midland Basin, are rank ordered: 1, rare (occurs in <10% of sampling quadrats at any given locality), 2, common (occurs in 10 to 50% of sampling quadrats), and 3, abundant (occurs in >50% of sampling quadrats). (A) Tree ferns and pteridosperms are hygromorphic and occur in deposits with sedimentologic and pedogenic indicators of everwet to subhumid seasonal conditions. Red climate curve for paleoequatorial western Euramerica defined by using soil moisture regimes and degree of seasonality inferred from paleosol morphologies (*23*); zigzag pattern indicates short-term (10^3 to 10^5 year) climate cycles inferred from intervals of polygenetic soils that exhibit climatically out-of-phase superposition of calcic and argillic horizons. (B) Conifers and peltasperms are xeromorphic and typically are found in association with sedimentologic and pedogenic indicators of moisture limitation.

The amplitudes of the reconstructed SST shifts (*40*) indicate substantial changes in the mean state of tropical climate during the end of the LPIA, with glacial tropical oceans at least 4° to 7°C cooler than those of intervening glacial minima (Fig. 2B). Inferred periods of elevated tropical SSTs and pCO_2 coincide with independently recognized intervals of warmer temperate conditions in high-latitude southern Gondwana (Fig. 2C) indicated by the accumulation of nonglacial sediments, including extensive kaolin and bauxite deposits in Australia during peak (Artinskian) warming and pCO_2 (*7*) and increased faunal diversity in Australia and South America (*7*, *11*, *41*). The covariance among inferred shifts in paleotropical SSTs, pCO_2, and variations in high-latitude Gondwanan glaciation and climate implies a strong CO_2-climate-glaciation linkage during the Permian. Although our coupled records suggest atmospheric CO_2 may have played a direct role in forcing Early to Middle Permian climate and ice mass stability, a determination of phase relationships between these parameters is precluded by the uncertainties in the age models. The inferred variations in tropical SSTs between periods of glacial maxima and minima, however, are consistent with the range predicted by Permian climate simulations for a change in radiative CO_2 forcing from 1 to 8 PAL (*42*).

Permo-Carboniferous plant assemblages from western paleoequatorial Euramerica archive a mechanistic vegetational response to late Paleozoic pCO_2 and climate change. Reconstructed plant communities from the same terrestrial successions that host the pedogenic mineral-bearing paleosols document major dominance-diversity changes corresponding one-for-one to inferred changes in paleotropical climate, pCO_2, and glacial extent (Fig. 3 and table S6). Four tropical biomes appear in succession, composed of increasingly xeromorphic species, representing progressively more seasonally moisture-stressed environments. These biomes are floristically distinct, sharing only a few opportunistic ferns and sphenopsids (*43*). Typical latest Carboniferous flora, rich in marattialean ferns, medullosan pteridosperms, sphenopsids, and sigillarian lycopsids, was replaced essentially instantaneously by one rich in conifers {*Walchia* and *Ernestiodendron*; compare with *Brachyphyllum* (*44*), callipterids (*Rhachiphyllum*), cycadophytes (*Russellites*), and other seed plants [*Cordaites*, *Sphenopteridium* (*45*)]}. This floristic shift is synchronous with an abrupt continental climate transition from everwet to semi-arid conditions (Fig. 3A), characterized by increased temperatures (*18*, *24*) and seasonal moisture availability inferred from paleosol morphologies (*23*).

Conifers and callipterids diversified in seasonally dry habitats during the initial Early Permian (Sakmarian) rise in CO_2 and the warm period of glacial minima, spatially replacing the tree fern–rich and the pteridosperm-rich wetland floras (Fig. 3). Tree fern–rich floras reappeared during wetter, cooler conditions of the mid-Early

Permian (Artinskian) glaciation, stratigraphically intercalated but not mixed, with conifer-callipterid floras. These two glacial floras show limited species overlap and oscillated at the 10^3- to 10^5-year scale, reflecting short-lived pluvials (*46*). Dramatic floristic changes also occurred during the cold period at the close of the Early Permian (Kungurian), with the migration into lowland basins of unique seed-plant assemblages not observed again until the Late Permian (conifers) and the Mesozoic (cycads) (*47*). These temporally successive floras tracked climatic conditions and contained progressively more evolutionarily advanced lineages. This suggests that evolutionary innovation, the appearance of new plant body plans, occurred in extrabasinal areas and was revealed by climate-driven floral migration into lowland basins.

The history of latest Carboniferous to Middle Permian climate provides a unique deep-time perspective on the precarious balance between icehouse and greenhouse states during major climate transitions, which are coupled to changing atmospheric CO_2 content. Maximum expansion of Gondwanan continental ice sheets occurred during earliest Permian time (*10*) under the lowest paleoatmospheric CO_2 levels and paleotropical SSTs. Widespread Early Permian (mid-Sakmarian) collapse of ice sheets (*8*, *10*) coincided with the onset of rising atmospheric CO_2 levels, after which time tropical SSTs and pCO_2 rose. Subsequent glacial influence was restricted to eastern Australia (*6*), with resurgent ice masses occurring during three more episodes (*11*) of lowered atmospheric pCO_2 before the permanent transition to an ice-free world (260 Ma). Our study indicates that ice buildup in Australia during subsequent cold periods, however, was progressively less widespread, with the two youngest glacials generally confined to local valleys or mountain ice caps along the polar margin of Australian Gondwana. Notably, SSTs and pCO_2 did not return to earliest Permian levels during these post-Sakmarian glacial periods.

Our reconstructed pCO_2, paleotemperatures, and inferred glacial history depict an Early Permian atmosphere that systematically increased from PAL to levels similar to those predicted to exist if fossil fuels are exhausted. Although global-scale deglaciation was unrelenting under rising Early Permian atmospheric CO_2, transient periods of icehouse stability and glacial resurgence returned during short-lived intervals of low pCO_2, perhaps until a CO_2 threshold and greenhouse stability precluded the reestablishment of glacial conditions [compare with (*48*)]. This late Paleozoic climate behavior mimics, in reverse, the magnitude and temporal scale of atmospheric CO_2 changes and ephemeral warmings that foreshadowed the transition into our present glacial state (*4*, *5*), further documenting the degree of climate variability, carbon cycle perturbation, and tropical ecosystem restructuring that has been associated with past CO_2-forced climate transitions.

References and Notes

1. J. R. Petit *et al.*, *Nature* **399**, 429 (1999).
2. U. Siegenthaler *et al.*, *Science* **310**, 1313 (2005).
3. L. R. Kump, *Nature* **419**, 188 (2002).
4. M. Pagani, J. C. Zachos, K. H. Freeman, B. Tipple, S. Bohaty, *Science* **309**, 600 (2005); published online 16 June 2005 (10.1126/science.1110063).
5. A. Tripati, J. Backman, H. Elderfield, P. Ferretti, *Nature* **436**, 341 (2005).
6. L. A. Frakes, J. E. Francis, J. I. Syktus, *Climate Modes of the Phanerozoic* (Cambridge Univ. Press, Cambridge, 1992).
7. J. M. Dickins, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **125**, 185 (1996).
8. H. Stollhofen, I. G. Stanistreet, B. Bangert, H. Grill, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **161**, 127 (2000).
9. C. R. Gonzalez, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **79**, 275 (1990).
10. J. L. Isbell, M. F. Miller, K. L. Wolfe, P. A. Lenaker, *Geol. Soc. Am. Spec. Pap. 370* (2003), pp. 5–24.
11. A. T. Jones, C. R. Fielding, *Geology* **32**, 153 (2004).
12. T. E. Cerling, in *Palaeoweathering, Palaeosurfaces, and Related Continental Deposits*, M. Thiry, R. Simon-Coinçon, Eds. (Blackwell, Cambridge, 1999), pp. 43–60.
13. C. J. Yapp, H. Poths, *Nature* **355**, 342 (1992).
14. D. L. Royer, R. A. Berner, D. J. Beerling, *Earth Sci. Rev.* **54**, 349 (2001).
15. D. D. Ekart, T. E. Cerling, I. P. Montañez, N. J. Tabor, D. J. Beerling, *Am. J. Sci.* **299**, 805 (1999).
16. CO_2 paleobarometer input parameters include, from most to least sensitive: (i) soil-respired CO_2 content (S_z in ppmv), typically assumed to be constant, (ii) the C isotopic composition of soil-respired ($\delta^{13}C_O$) and atmospheric ($\delta^{13}C_A$) CO_2, both typically derived from broadly contemporaneous marine carbonates, and (iii) temperature, typically held constant (~25°C). Pedogenic carbonate $\delta^{13}C$ is utilized as the proxy of $\delta^{13}C$ of total soil-CO_2. For goethite-based pCO_2 estimates, the measured concentration and $\delta^{13}C$ of the $Fe(CO_3)OH$ component in pedogenic goethite are taken as proxies of ambient pCO_2 and $\delta^{13}C$, respectively, present during crystallization; $\delta^{13}C_O$ and $\delta^{13}C_A$ are inferred from marine carbonate proxies or fossil organic matter.
17. Materials and methods are available as supporting material on *Science* Online.
18. N. J. Tabor, I. P. Montañez, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **223**, 127 (2005).
19. The $\delta^{13}C_{org}$ values of matrix in calcic-paleosols and contemporaneous fossil-plant matter overlap (Fig. 1B), indicating that the latter are representative of soil-respired CO_2 in carbonate-bearing paleosols. A $\delta^{13}C_{atm}$ proxy record was calculated assuming isotopic equilibrium fractionation ($\Delta_{atm\text{-}om}$ of 18.5‰) between paleoflora ($\delta^{13}C_{org}$) and atmospheric CO_2. Paleo-pCO_2 estimated using a 2‰ variation in $\Delta_{atm\text{-}om}$ falls within the uncertainty band shown on Fig. 1C, reflecting the low sensitivity of the model to $\delta^{13}C_A$.
20. N. C. Arens, A. H. Jahren, R. Amundson, *Paleobiology* **26**, 137 (2000).
21. K. M. Panchuk, C. Holmden, S. A. Leslie, *J. Sediment. Res.* **76**, 200 (2006).
22. H. Zhang, G. Shen, Z. He, *Acta Geol. Sin.* **73**, 111 (1999).
23. N. J. Tabor, I. P. Montañez, *Sedimentology* **51**, 851 (2004).
24. N. J. Tabor, *Earth Planet. Sci. Lett.*, in press.
25. N. J. Tabor, C. J. Yapp, I. P. Montañez, *Geochim. Cosmochim. Acta* **68**, 1503 (2004).
26. M. R. Saltzman, *Geology* **31**, 151 (2003).
27. K. Scheffler, S. Hoernes, L. Schwark, *Geology* **31**, 605 (2003).
28. R. A. Berner, Z. Kothavala, *Am. J. Sci.* **301**, 182 (2001).
29. C. I. Mora, S. G. Driese, L. A. Colarusso, *Science* **271**, 1105 (1996).
30. Published $\delta^{18}O$ data are from biostratigraphically constrained and diagenetically screened calcitic brachiopods (*31*, *35*, *49*); $\delta^{18}O$ values < −8‰ were excluded. The $\delta^{18}O$ values were detrended to remove the effects of the long-term linear Phanerozoic trend by removing, from each data point, the least squares linear fit calculated using SPlus.
31. H. Mii, E. Grossman, T. E. Yancey, *Geol. Soc. Am. Bull.* **111**, 960 (1999).
32. B. Bangert, R. Armstrong, H. Stollhofen, V. Lorenz, *J. Afr. Earth Sci.* **29**, 33 (1999).
33. M. R. Banks, in *Earth's Pre-Pleistocene Glacial Records*, M. J. Hambrey, W. B. Harland, Eds. (Cambridge Univ. Press, Cambridge, 1981), pp. 495–501.
34. J. N. J. Visser, *Sedimentology* **44**, 507 (1997).
35. C. Korte, T. Jasper, H. W. Kozur, J. Veizer, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **224**, 333 (2005).
36. D. P. Schrag *et al.*, *Quat. Sci. Rev.* **21**, 331 (2002).
37. C. B. Cecil *et al.*, in *Climate Controls on Stratigraphy*, C. B. Cecil, T. N. Edgar, Eds. (SEPM Special Publication 77, Society for Sedimentary Geology, Tulsa, OK, 2003), pp. 151–180.
38. D. L. Royer, R. A. Berner, I. P. Montañez, N. J. Tabor, D. J. Beerling, *Geol. Soc. Am. Today* **14**, 3 (2004).
39. W. C. Beck, E. L. Grossman, J. W. Morse, *Geochim. Cosmochim. Acta* **69**, 3493 (2005).
40. Detrended $\delta^{18}O$ values were binned into 1- to 3-My windows. Average values (±2 SE) were translated to a range of paleotropical SSTs by using a quadratic equation (*50*) and a constant (0‰) or variable (−0.5 to +1.5‰) $\delta^{18}O_{sw}$. The influence of local and environmental factors and diagenesis on the long-term $\delta^{18}O_{carb}$ trend is considered secondary, given the overlap in individual published datasets and Tethyan and Panthalassan brachiopod $\delta^{18}O$ values; secondary influences are likely recorded in the $\delta^{18}O_{brachiopod}$ variability within contemporaneous populations.
41. O. R. Lopez-Gamundi, in *Late Glacial and Postglacial Environmental Changes: Quaternary, Carboniferous-Permian, and Proterozoic*, I. P. Martini, Ed. (Oxford Univ. Press, Oxford, 1997), pp. 147–168.
42. A. M. E. Winguth *et al.*, *Paleoceanography* **17**, 1057 (2002).
43. W. A. DiMichele, R. B. Aronson, *Evolution* **46**, 807 (1992).
44. Compare with *Brachyphyllum* after S. H. Mamay, *U.S. Geol. Surv. Prof. Pap. 575-C* (1967), p. 120.
45. *Sphenopteridium*, after S. H. Mamay, *Am. J. Bot.* **79**, 1092 (1992).
46. W. A. DiMichele, N. J. Tabor, D. S. Chaney, W. J. Nelson, *Geol. Soc. Am. Spec. Pub. 299* (2006), pp. 223–248.
47. W. A. DiMichele, D. S. Chaney, W. H. Dixon, W. J. Nelson, R. W. Hook, *Palaios* **15**, 524 (2000).
48. L. R. Kump, *Nature* **436**, 333 (2005).
49. J. Veizer *et al.*, *Chem. Geol.* **161**, 59 (1999).
50. S. T. Kim, J. R. O'Neil, *Geochim. Cosmochim. Acta* **61**, 3461 (1997).
51. A chronostratigraphic framework for the Eastern Shelf of Texas (*23*) was used as the basis for the composite section into which other records were integrated with use of available biostratigraphy and lithostratigraphic correlation. Calibration of the Eastern Shelf succession to the Gradstein *et al.* 2004 geologic time scale (*52*) is based, for the study area, on the stratigraphic position of the stage boundaries determined by conodont, fusulinid, and ammonite biostratigraphy.
52. F. M. Gradstein, J. G. Ogg, A. G. Smith, Eds., *A Geologic Time Scale 2004* (Cambridge Univ. Press, Cambridge, 2005).
53. C. A. Ross, J. R. P. Ross, in *Sea-Level Changes: An Integrated Approach*, C. K. Wilgus *et al.*, Eds. (SEPM Special Publication 42, Society for Sedimentary Geology, Tulsa, OK, 1988), pp. 227–247.

54. We thank D. Chaney and G. Mack for field guidance and discussions; J. Fong for the illustrations; and H. Jahren, R. Berner, T. Cerling, and three anonymous reviewers for critical insight. Supported by NSF grants EAR-9914640, EAR-0417578, EAR-0545701, OPP-0126086, and ANT-0440919.

Supporting Online Material

www.sciencemag.org/cgi/content/full/315/5808/87/DC1
Materials and Methods
Fig. S1
Tables S1 to S6

22 August 2006; accepted 20 November 2006
10.1126/science.1134207

Late-Neoproterozoic Deep-Ocean Oxygenation and the Rise of Animal Life

Don E. Canfield,[1]* Simon W. Poulton,[2] Guy M. Narbonne[3]

Because animals require oxygen, an increase in late-Neoproterozoic oxygen concentrations has been suggested as a stimulus for their evolution. The iron content of deep-sea sediments shows that the deep ocean was anoxic and ferruginous before and during the Gaskiers glaciation 580 million years ago and that it became oxic afterward. The first known members of the Ediacara biota arose shortly after the Gaskiers glaciation, suggesting a causal link between their evolution and this oxygenation event. A prolonged stable oxic environment may have permitted the emergence of bilateral motile animals some 25 million years later.

Large, architecturally complex life forms first appeared about 575 million years ago (Ma) (*1*, *2*). These life forms are epitomized by the Ediacara biota, a globally distributed assemblage of fossil impressions of centimeter- to meter-scale soft-bodied organisms and colonies. The Ediacara biota developed after the Gaskiers glaciation at 580 Ma and abruptly disappeared coincident with the Cambrian explosion of skeletal animals about 35 million years later (*3*, *4*). The Ediacara biota most likely included a mixture of stem-group animals and potentially other extant kingdoms of eukaryotes, along with fossils that may represent "failed experiments" in evolution (*3*). Fossilized animal embryos (*5*) dated at 560 to 580 Ma (*6*) further support the view that animals were an important part of Ediacaran life. Animals have an absolute requirement for oxygen, and it has been widely hypothesized that a late-Neoproterozoic rise in oxygen created an environment permissive for animal evolution (*7*–*10*). Direct evidence for late-Neoproterozoic ocean oxygenation, however, has been lacking. We used several geochemical tracers to show that the deep ocean became oxidized shortly before the first appearance of the Ediacara biota.

In the Avalon Peninsula, Newfoundland, 6 km of late-Neoproterozoic sedimentary rocks represent at least 15 million years of late-Neoproterozoic time (Fig. 1). The sediment sequence begins about 800 m below the glacial deposits and cap carbonate of the Gaskiers Formation dated at 580 Ma (*2*). The Gaskiers glaciation was the last of the major Neoproterozoic glaciations (*4*, *11*) and is represented on four continents (*12*), although it was probably not as widespread as the previous Sturtian and Marinoan "snowball" glaciations (*11*, *12*). Ediacara-type fossils are present through most of the post-Gaskiers succession. The first Ediacaran fossils, represented by a low-diversity assemblage of *Ivesheadia* ("pizza disks"), *Thectardis*, and species of *Charnia* (*1*, *3*, *4*, *13*) (Fig. 2, A and B), were found in the upper Drook Formation within 5 million years of the glacial terminus. These are the earliest large and architecturally complex eukaryote fossils known anywhere in the world. By 565 Ma, the high-density and high-diversity Ediacaran assemblages of Mistaken Point existed (Fig. 2C) (*3*, *13*, *14*), and numerous occurrences of Ediacaran disks continue through the Fermeuse Formation (Fig. 2D).

The Ediacaran animals of the Avalon Peninsula lived on the sea floor in deep-water environments well below the photic zone (*3*, *13*, *15*). The Mall Bay through Briscal Formations (Fig. 1) accumulated as axial basin floor deposits in water depths exceeding several hundred meters, and possibly extending 1 km or more (*3*, *16*). The Mistaken Point and Trepassey Formations represent a deep-water slope environment (*14*, *15*). Beginning with the Fermeuse Formation, the sequence gives way to mudstones likely deposited as delta front deposits (*3*, *17*). The Neoproterozoic sediments of Newfoundland were deposited at the northern margins of ancient Gondwana, with no evidence for basin restriction, suggesting open access to the global ocean (*18*, *19*).

We used iron extraction techniques to explore ocean redox conditions. In this approach, operationally defined iron extraction protocols (*20*) are used to partition iron into its highly reactive components and its unreactive phases. Highly reactive iron includes iron oxide, carbonate, and sulfide minerals; this represents the iron that is geochemically and biologically active during early sediment diagenesis (*21*). By contrast, "unreactive" iron is geochemically inert on early diagenetic time scales. Previous studies have shown that in a broad suite of marine sediments deposited from an oxygen-containing water column, the ratio of highly reactive iron to total iron (FeHR/FeT) is consistently below 0.38, with a modern average (±SD) of 0.26 ± 0.08 (*22*, *23*) and a Phanerozoic average (past 542 million years and excluding the modern) of 0.15 ± 0.06 (*23*). By contrast, sediments deposited from anoxic water columns may obtain additional reactive iron from iron mineral formation in the water column; in these environments, FeHR/FeT may exceed 0.38.

This is true both if the anoxic water column is sulfidic, such as the modern Black Sea (*22*) and ancient sulfidic marine water bodies (*23*, *24*), and if it contained dissolved iron, as was the case early in Earth's history (*25*). We also used the concentrations of organic carbon and pyrite sulfur, as well as the isotopic composition of sulfur, to aid our characterization of the depositional environment (*26*) (table S1).

There is a marked difference in the proportion of highly reactive iron in sediments depos-

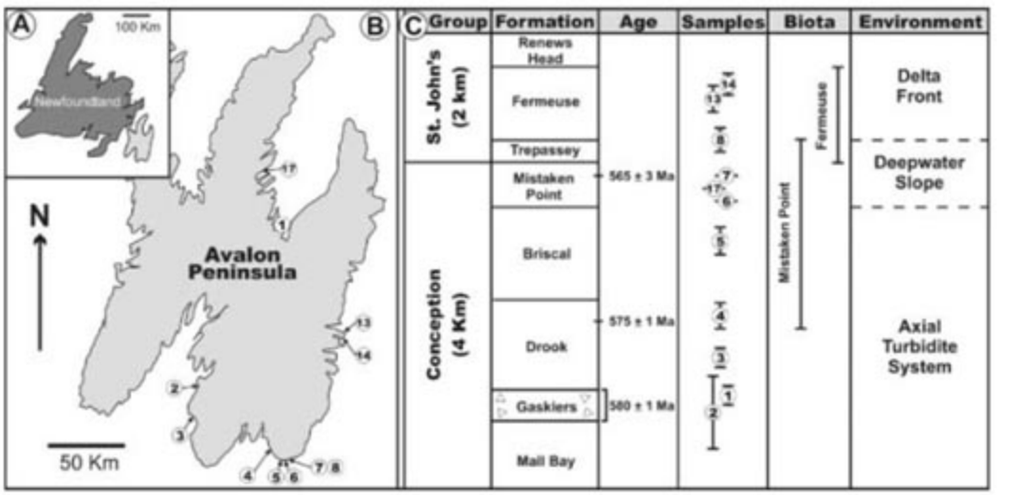


Fig. 1. Location (**A** and **B**) and stratigraphic setting (**C**) of geochemical samples. Location numbers are the same as in (*19*) (sites 1, 2, and 4 to 17) and (*39*) (sites 1 to 8, 13, and 14), which contain details on the precise location, access, lithostratigraphy, sedimentology, and fossil taxa for each locality. The supporting online material includes stratigraphic positioning of the samples (*26*). U/Pb dates are from (*2*, *40*).

[1]Nordic Center for Earth Evolution (NordCEE) and Institute of Biology, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark. [2]School of Civil Engineering and Geosciences, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK. [3]Department of Geological Sciences and Geological Engineering, Queen's University, Kingston, Ontario K7L 3N6, Canada.

*To whom correspondence should be addressed. E-mail: dec@biology.sdu.dk

ited before and after the Gaskiers glaciation (Fig. 3A). In numerous instances, the Gaskiers diamictite shows FeHR/FeT ratios exceeding 0.38, indicative of anoxic deposition. This is also the case, to a lesser extent, in the upper Mall Bay Formation. In other cases, the FeHR/FeT ratio of Gaskiers and Mall Bay samples is less than 0.38 but more than both the modern and Phanerozoic average ratios (Fig. 3A). In these cases, the extraction results are somewhat equivocal, but given the numerous instances in which FeHR/FeT exceeds 0.38 and the high overall ratios of FeHR/FeT, the results suggest that the water column was anoxic during deposition of the Gaskiers diamictite and probably also the upper Mall Bay (*27*). There is little sulfide sulfur in these rocks, and most of the reactive iron is bound as iron oxide and iron carbonate (*26*). Therefore, a sulfidic water column can be ruled out (*25*), indicating that the water column was most likely ferruginous, containing elevated concentrations of dissolved ferrous iron (Fe^{2+}). Therefore, the deep-ocean chemistry accompanying the Gaskiers glaciation (and possibly also the Mall Bay) appears similar to the water chemistry associated with the earlier "snowball Earth" glaciations, in which banded iron formations accompanied the deposition of Sturtian-aged diamictites (about 700 Ma) (*11*) and Fe-enriched carbonates were deposited immediately after the Marinoan glaciation (about 630 Ma) (*28*), both of which are indicative of ferruginous oceanic conditions.

In sediment accumulated immediately after the Gaskiers glaciation and during a period of time representing more than 15 million years, most FeHR/FeT ratios are much lower than in the underlying rocks and fall between the Phanerozoic and modern average for oxic sediment deposition; nearly all are less than 0.38 (Fig. 3A). These data provide evidence for a long period of stable deep-water oxic marine conditions. Overall, our data point to a pre-Drook anoxic iron-containing ocean giving way to oxic marine conditions after the Gaskiers glaciation.

For the Mall Bay through the Trepassey Formations, organic carbon concentrations are, in general, extremely low and are completely consistent with deposition in oligotrophic deep-water basinal or outer slope marine settings (*29*). Whereas higher organic carbon concentrations might be expected under anoxic conditions (*30*), our rather low concentrations in the Mall Bay and Gaskiers Formations are similar to those found (0.1 to 0.3 weight % C) in the Sturtian-aged (about 730 Ma) Rapitan Iron Formation, which also deposited from anoxic Fe-containing waters (*31*). The somewhat higher concentrations observed in the Fermeuse Formation are consistent with an environment closer to shore of higher sediment deposition rate such as a delta front would offer.

The isotopic composition of sulfur is quite variable but demonstrates patterns consistent with the above scenario for deep-ocean oxygenation, providing further insights into the nature of ocean chemistry. The isotopic composition of sulfide in pre-Gaskiers sediments is consistently greater than zero, indicating relatively small fractionations from seawater sulfate of around 18 ± 10 per mil (Fig. 3C) [the isotopic composition of seawater sulfate from 590 to 560 Ma is well constrained at between 22 and 28 per mil (*32*)]. This, combined with generally low concentrations of organic carbon and sulfide, is consistent with low rates of sulfate reduction under sulfate-limiting concentrations. Previous modeling (*33*) suggests that such low fractionations should occur with submillimolar sulfate concentrations. Higher fractionations in the Gaskiers, and particularly in the Drook Formation, demonstrate a change in the sulfur cycle. Indeed, the higher post–Mall Bay fractionations are consistent with an increase in sulfate concentration, which allows the expression of higher fractionations when compared with those produced with low sulfate levels. This pattern of increased fractionations also occurs with sediments deposited in association with and immediately after the Sturtian and Marinoan glaciations (*34*) and is thus a general feature of Neoproterozoic glacial and postglacial deposits.

To explain these increased fractionations, we suggest that glacial melting increased the nutrient load to the ocean. This stimulated primary production and carbon burial and thus increased atmospheric oxygen levels. Increased oxygen enhanced the oxidative weathering of sulfide to sulfate on the continents, thus increasing the flux of sulfate to the ocean and marine sulfate concentrations. The presence of anoxygenic photosynthetic biomarkers in post-Sturtian and post-Marinoan deposits argues against substantial ocean oxygenation after the earlier Neoproterozoic glaciations (*35*, *36*), but the post-Gaskiers event was sufficient to result in oxygenation of the deep ocean. A return to much lower fractionations in post–Mistaken Point sediments could reflect a return to lower marine sulfate concentrations or, perhaps more likely, a reduction in isotope fractionation as might be expected with higher rates of sediment deposition and increased rates of sulfate reduction in a delta-front environment (*33*).

Our evidence for deep-water oxygenation in the post-Gaskiers ocean may place some constraints on the minimum level of atmospheric oxygen at this time. We reason as follows: Organic matter produced in the surface ocean consumes oxygen during degradation as it falls through the water column. In the modern ocean, oxygen concentrations reach a minimum at depths of 500 to 1500 m (*37*). The magnitude of the oxygen deficit is about 40 to 100 μM in the North Atlantic and 100 to 300 μM in the North Pacific (*37*). If we assume that the ocean structure was similar to its structure today and that the Drook and Briscal Formations were deposited in water depths of 500 to 1500 m, then oxygen deficits would have ranged somewhere between 40 and 300 μM. If we take 40 μM as the most conservative estimate, then at least this much O_2 was dissolved in the waters supplying the deep ocean. Probably a bit more oxygen was required, given that the Ediacara biota would have likely needed 10 to 20 μM for their respiration (*38*), which we add to our minimum estimate of the oxygen content of the water supplying the deep ocean. At present,

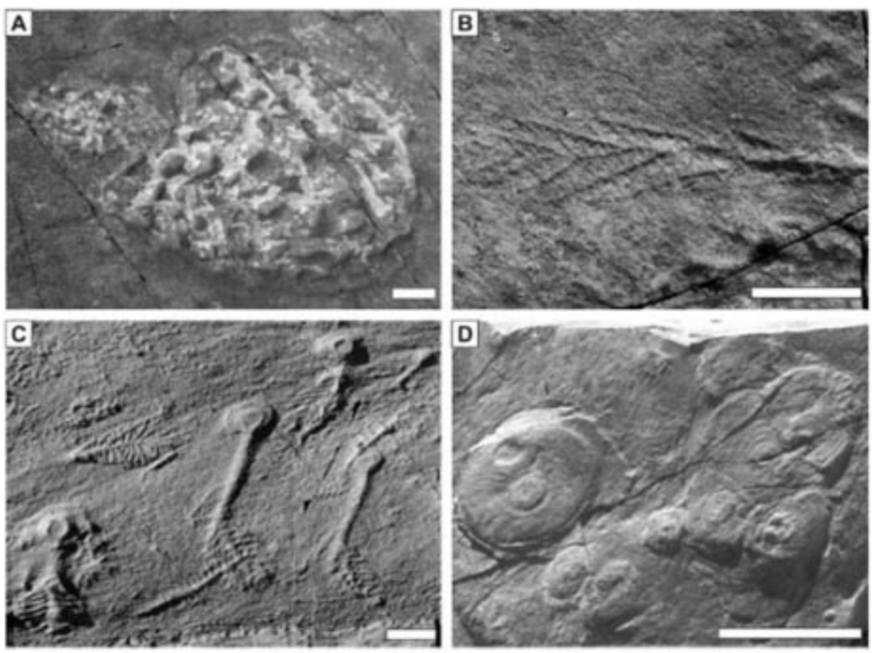


Fig. 2. Ediacaran fossils from Newfoundland. Scale bars represent 5 cm. (**A** and **B**) Low-diversity assemblage of fossils from the Drook Formation (575 Ma) at locality 4. (A) Two specimens of the discoid fossil *Ivesheadia*. (B) *Charnia* frond. (**C**) High-diversity Mistaken Point assemblage (565 Ma), exhibiting large fronds (*Charniodiscus*) along with frondose, bushlike, and spindle-shaped rangeomorph fossils at locality 7. (**D**) Fermeuse assemblage (about 560 Ma) of *Aspidella* disks at locality 13.

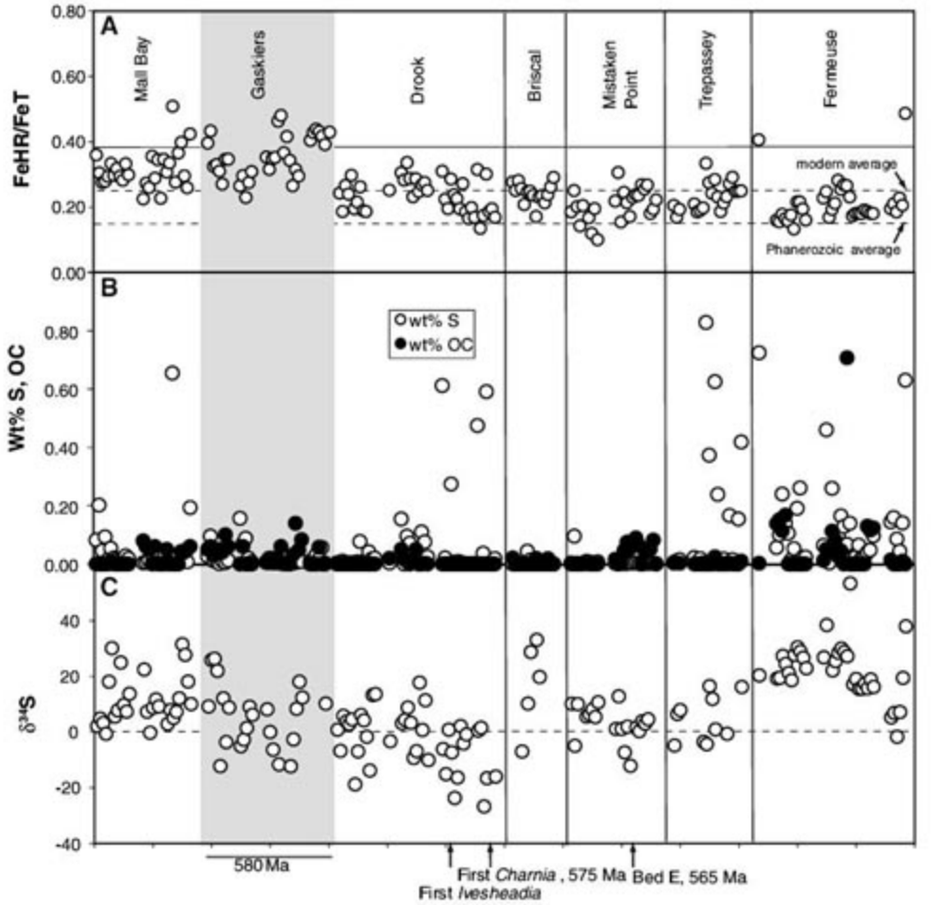


Fig. 3. The chemistry of late-Neoproterozoic sediments from the Avalon Peninsula, Newfoundland. (**A**) FeHR/FeT ratios. Solid line represents the 0.38 ratio. Dashed lines represent the ratios for average modern and average Phanerozoic sediments deposited in an oxic water column. (**B**) Concentrations of reduced sulfur (pyrite) and organic carbon (OC) by weight (wt) %. (**C**) Isotopic composition of pyrite sulfur. A dashed line at 0 per mil is shown for reference. Also indicated are key dates (as in Fig. 1) and the stratigraphic level of prominent fossil locations.

deep water is formed at high latitudes with an air-saturated O_2 concentration of 325 μM. If we require a minimum of 50 μM O_2 in this water, then we need to saturate with atmosphere containing greater than 15% of present day oxygen levels. Thus, 15% of present day oxygen levels is a minimum estimate for post-Gaskiers atmospheric O_2.

In the Avalon region, this oxygenation was stable and persisted for at least 15 million years. In our preferred scenario, this oxygenation was widespread, marking the first time that oxygen concentrations reached levels permissive for the metabolism of large multicellular heterotrophic eukaryotes. If so, evolution into this new permissive ecology could have been quite rapid (*7*), resulting in the emergence of the Ediacara biota within 5 million years (*1*, *2*), and recognizable motile animals within another 20 million years (*3*, *4*). In another scenario, the Ediacaran biota populating Avalonia evolved earlier, perhaps before the Gaskiers glaciation, and migrated to Avalon after oxygenation of the local environment. The available data support the first scenario, but further exploration of pre-Gaskiers ocean chemistry and biology will help to elucidate the possibility of the second.

References and Notes

1. G. M. Narbonne, J. G. Gehling, *Geology* **31**, 27 (2003).
2. S. A. Bowring, P. Myrow, E. Landing, J. Ramenzani, *Astrobiology* **2**, 112 (2002).
3. G. M. Narbonne, *Annu. Rev. Earth Planet. Sci.* **33**, 421 (2005).
4. A. H. Knoll, M. R. Walter, G. M. Narbonne, N. Christie-Blick, *Lethaia* **39**, 13 (2006).
5. S. Xiao, Y. Zhang, A. H. Knoll, *Nature* **391**, 553 (1998).
6. D. Condon, M. Zhu, S. Bowring, W. Wang, J. Yang, *Science* **308**, 95 (2005).
7. A. H. Knoll, *Life on a Young Planet: The First Three Billion Years of Evolution on Earth* (Princeton Univ. Press, Princeton, NJ, 2003).
8. L. V. Berkner, L. C. Marshall, *J. Atmos. Sci.* **22**, 225 (1965).
9. A. H. Knoll, *Science* **256**, 622 (1992).
10. C. R. Marshall, *Annu. Rev. Earth Planet. Sci.* **34**, 355 (2006).
11. P. F. Hoffman, D. P. Schrag, *Terra Nova* **14**, 129 (2002).
12. G. P. Halverson, in *Neoproterozoic Geobiology and Paleobiology, Topic in Geobiology*, S. Xiao, A. J. Kaufman, Eds. (Kluwer, New York, 2005), vol. 27, pp. 231–271.
13. M. E. Clapham, G. M. Narbonne, J. G. Gehling, *Paleobiology* **29**, 527 (2003).
14. S. B. Misra, *Geol. Soc. Am. Bull.* **82**, 979 (1971).
15. D. A. Wood, R. W. Dalrymple, G. M. Narbonne, J. G. Gehling, M. E. Clapham, *Can. J. Earth Sci.* **40**, 1375 (2003).
16. S. Gardiner, R. N. Hiscott, *Can. J. Earth Sci.* **25**, 1579 (1988).
17. H. Williams, A. F. King, *Geol. Surv. Can. Mem.* **389** (1979).
18. C. R. Scotese, PALEOMAP Project (Arlington, TX, 2001), vol. 51.
19. G. M. Narbonne, R. W. Dalrymple, M. La Flamme, J. Gehling, W. D. Boyce, *Life After Snowball: The Mistaken Point Biota and the Cambrian of Newfoundland, NAPC Field Trip Guide* (North American Paleontological Convention, Halifax, Canada, 2005).
20. S. W. Poulton, D. E. Canfield, *Chem. Geol.* **214**, 209 (2005).
21. D. E. Canfield, *Geochim. Cosmochim. Acta* **53**, 619 (1989).
22. R. Raiswell, D. E. Canfield, *Am. J. Sci.* **298**, 219 (1998).
23. S. W. Poulton, R. Raiswell, *Am. J. Sci.* **302**, 774 (2002).
24. Y. Shen, D. E. Canfield, A. H. Knoll, *Am. J. Sci.* **302**, 81 (2002).
25. S. W. Poulton, P. Fralick, D. E. Canfield, *Nature* **431**, 173 (2004).
26. Materials and methods are available as supporting material on *Science* Online.
27. Although FeHR/FeT ratios exceeding 0.38 are strongly indicative of sediment deposition in an anoxic water column, anoxic deposition can also yield lower ratios if the flux of clastic material is sufficient to overwhelm the water-column flux of iron. This happens, for example, in turbidite intervals in the Black Sea (*22*). Also, localized situations of preferential continentally derived iron oxide deposition or sediment iron remobilization can yield FeHR/FeT ratios of greater than 0.38, as is sometimes the case in inner shore environments (*23*). However, we are aware of no instances whereby oxic deposition in fully marine settings has resulted in FeHR/FeT ratios of greater than 0.38, and this is the case for both continental margin sediments and deep-sea sediments such as red clays (*23*). As further evidence of enhanced iron deposition in pre-Drook sediments, we highlight that many intervals of the Gaskiers diamictite weather blood-red in the field, in stark contrast to the greenish gray color of the Drook.
28. M. T. Hurtgen, G. P. Halverson, M. A. Arthur, P. F. Hoffmann, *Earth Planet. Sci. Lett.* **245**, 551 (2006).
29. E. T. Premuzic, C. M. Benkovitz, J. S. Gaffney, J. J. Walsh, *Org. Geochem.* **4**, 63 (1982).
30. D. E. Canfield, *Chem. Geol.* **114**, 315 (1994).
31. C. Klein, N. J. Beukes, *Econ. Geol.* **88**, 542 (1993).
32. D. A. Fike, J. P. Grotzinger, L. M. Pratt, R. E. Summons, *Nature* **444**, 744 (2006).
33. K. S. Habicht, M. Gade, B. Thamdrup, P. Berg, D. E. Canfield, *Science* **298**, 2372 (2002).
34. P. Gorjan, J. J. Veevers, M. R. Walter, *Precambrian Res.* **100**, 151 (2000).
35. A. N. Olcott, A. L. Sessions, F. A. Corsetti, A. J. Kaufman, T. Flavio de Oliviera, *Science* **310**, 471 (2005).
36. A. N. Olcott, C. Li, A. L. Sessions, F. A. Corsetti, P. Peng, *Geochim. Cosmochim. Acta* **70**, A456 (2006).
37. W. S. Broecker, T.-H. Peng, *Tracers in the Sea* (Lamont-Doherty Geological Observatory, Palisades, NY, 1982).
38. B. Runnegar, *Global Planet. Change* **5**, 97 (1991).
39. G. M. Narbonne, R. W. Dalrymple, J. G. Gehling, *Neoproterozoic Fossils and Environments of the Avalon Peninsula, Newfoundland*. (Geological Association of Canada, Guidebook B5, St. John's, Canada, 2001).
40. A. P. Benus, in *Trace Fossils, Small Shelly Fossils and the Precambrian- Cambrian Boundary*, E. Landing, G. M. Narbonne, P. Myrow, Eds. (New York State Museum and Geological Survey Bulletin, Albany, NY, 1988), vol. 463, pp. 8–9.

41. Our work was funded by the Danish National Research Foundation (Danmarks Grundforskningsfond), a Natural Environment Engineering Research Council Research Fellowship (S.W.P.), the Natural Sciences and Engineering Research Council of Canada, and a Queen's Research Chair. We thank the Parks and Natural Areas, and Government of Newfoundland and Labrador for scientific permits; the Portugal Cove South Interpretation Centre for help and support; and M. Laflamme and L. Salling for help.

Supporting Online Material
www.sciencemag.org/cgi/content/full/1135013/DC1
Materials and Methods
Table S1
References

12 September 2006; accepted 27 October 2006
Published online 7 December 2006;
10.1126/science.1135013
Include this information when citing this paper.

Climate Change Affects Marine Fishes Through the Oxygen Limitation of Thermal Tolerance

Hans O. Pörtner* and Rainer Knust

A cause-and-effect understanding of climate influences on ecosystems requires evaluation of thermal limits of member species and of their ability to cope with changing temperatures. Laboratory data available for marine fish and invertebrates from various climatic regions led to the hypothesis that, as a unifying principle, a mismatch between the demand for oxygen and the capacity of oxygen supply to tissues is the first mechanism to restrict whole-animal tolerance to thermal extremes. We show in the eelpout, *Zoarces viviparus*, a bioindicator fish species for environmental monitoring from North and Baltic Seas (Helcom), that thermally limited oxygen delivery closely matches environmental temperatures beyond which growth performance and abundance decrease. Decrements in aerobic performance in warming seas will thus be the first process to cause extinction or relocation to cooler waters.

Climate change is projected to affect individual organisms, the size and structure of their populations, the species composition of communities, and the structure and functioning of ecosystems. Effects include poleward or high-altitude shifts in the distribution of ectothermic animals (*1*). A comprehensive mechanistic understanding has so far been lacking (*2*) but is needed for prediction of climate change effects. Physiological studies can address the mechanisms and reasons for the thermal sensitivity of organisms and their life stages.

In aquatic animals, a decrease in the capacity to perform aerobically (a drop in aerobic scope) characterizes the onset of thermal limitation at both ends of the thermal envelope [pejus thresholds T_p, fig. S1 (*3*–*6*)]. The reduction in aerobic scope is caused by limited capacity of circulatory and ventilatory systems to match oxygen demand. Such a constraint affects all higher functions (muscular activity, behavior, growth, and reproduction) and might thereby shape the long-term fate of species. Aerobic scope becomes minimal beyond low or high critical temperatures (T_c). Survival is then passive and time-limited, supported by anaerobic metabolism and protection of proteins and membranes by heat shock proteins and antioxidative defense. Thermal tolerance is hierarchical, with narrowing windows from molecular to cellular to systemic levels (*6*). Temperate species are able to acclimatize and shift the thermal window through changes in mitochondrial densities as well as other molecular to systemic adjustments of functional capacities (*3*, *6*–*10*) (fig. S1). Limits to acclimatization are set by trade-offs at various structural and functional levels that constrain the width of the thermal window, for example, through the trend to minimize energy turnover in relation to climate variability (*9*, *10*).

We investigated thermal limitation of the common eelpout, *Zoarces viviparus*, in its southernmost distribution area, the German Wadden Sea (part of the southern North Sea) during summer and thereby tested the ecological relevance of the concept of oxygen- and capacity-limited thermal tolerance (fig. S1). During the past 40 years, water temperatures in the German Bight increased by 1.13°C (at Helgoland Roads). Cold winters with sea surface temperatures (SSTs) around –1°C had occurred about once every 10 years up to 1944 but were experienced only once since 1960 (*11*). Models predict further SST increments for the next 90 to 100 years, by about 1.6° to 3.0°C in the northern and even by 3.0° to 3.9°C in the shallower southern North Sea (*12*), accompanied by rising sea levels (13 to 68 cm by 2050) and an increasing frequency of storm events (*13*).

Comparison of existing data sets indicates that field observations can be explained by the eelpout's physiological responses to warming (Fig. 1). The relative abundance of the nonmigratory eelpout decreases upon warming (5-year running means, data from 1954 to 1989, Fig. 1A) (*14*, *15*), reflecting a higher mortality in hot summers. Reduced field abundance coincides with reduced growth of laboratory-maintained, temperature-acclimated individuals (Fig. 1B). Individual growth is a key parameter shaping population growth and depends on aerobic scope. Lopsided growth curves result from the exponential rise in net aerobic scope upon warming, which is counterbalanced by the concomitant exponential rise in baseline metabolic costs (Fig. 1). Both abundance and growth begin to fall beyond upper pejus temperatures (T_p) (Fig. 1, C to E), reflecting the species-specific limits of acclimation capacity.

Pejus temperatures were derived from limitations in circulatory capacity (Fig. 1C), which occur before ventilatory limitations in eelpout (*Z. viviparus* and *Pachycara brachycephalum*) and Atlantic cod (*Gadus morhua*) (*4*, *16*–*18*). The loss of aerobic scope can also be derived from the shift of critical oxygen tensions, P_c, or concentrations, $[O_2]_c$. P_c or $[O_2]_c$ indicate oxygen limitation to the passive organism in hypoxia and the onset of anaerobic metabolism. Upon warming, $[O_2]_c$ reaches air saturation at T_c, where anaerobic metabolism begins in animals exposed to fully aerated waters (Fig. 1D). Aerobic scope thus begins to fall when $[O_2]_c$ starts to rise beyond T_p (Fig. 1E). Warming exacerbates oxygen limitations not only by the forced rise in oxygen demand, but also by reducing oxygen solubility (Fig. 1E).

The analysis of ecological responses in relation to 5-year running means of summer maxima, albeit improving the signal-to-noise ratio, may not precisely quantify temperatures and mechanisms effective in the field. Analysis of individual summers in long-term data series (*19*) should provide more detailed insight into cause-and-effect relationships (Figs. 2 and 3). The limited data set indicates that extreme temperatures of previous summers cause reduced abundance. Sampling took place in July, so the effects of the hottest season only become visible in the next year. Thermal limitation of aerobic scope may also translate into the next year by reducing the degree of successful fertilization and reproduction.

Thermal sensitivity is likely to be enhanced at large body sizes. In contrast to eelpout from the Baltic or from colder regions like the Russian White Sea, eelpout of the Wadden Sea only reach a maximum body length of about 23 cm at a maximum age of 3 to 4 years (*20*). A preliminary analysis of seasonal changes in size frequency distribution (fig. S2) shows that older specimens (larger than 20 cm) have low overall abundance and thus high mortality rates. High

Alfred Wegener Institute for Polar and Marine Research, Animal Ecophysiology, Postfach 12 01 61, 27515 Bremerhaven, Germany.

*To whom correspondence should be addressed. E-mail: hpoertner@awi-bremerhaven.de

mortality of large fish is probably not associated with predation, which usually occurs during early life stages (*21*). Rather, the oxygen-limitation model predicts that temperature-dependent aerobic limits are experienced earlier by larger than by smaller individuals (*22*). In fact, thermal sensitivity of growth or exercise was found to be enhanced in large compared with small individuals of various fish species (*23*, *24*). Thermal limits depicted in Figs. 1 to 3 are valid for specimens of about 23 cm body length (compare with fig. S2). Mild summers, with temperatures regularly beyond T_p of large fish, can therefore be interpreted to cause mortality of this size class (Fig. 3 and fig. S2). Because of wider thermal windows in smaller specimens, these temperatures still allow for population growth, seen especially during the mild summer of 1998 (Fig. 2). The earlier loss in aerobic scopes of large individuals indicates that specimens do not grow beyond oxygen-dependent size limits set by temperature such that this size group displays low abundance all year round (fig. S2). The species finally experiences a net reduction in abundance (Fig. 1A) when smaller individuals are also affected and population loss during hot summers exceeds yearly population growth. In conclusion, harmful effects of warming set in beyond pejus temperatures. Only summers hotter than critical temperatures of the larger specimens (Figs. 1 to 3) entail the full range of thermal stress phenomena depicted in fig. S1. The mismatch in oxygen supply versus demand thus becomes effective at the ecosystem level before the onset of anaerobic metabolism or of thermal damage (Td in fig. S1) and also before critical thermal maxima (CT_{max}) traditionally determined in thermal biology (*25*).

Overall, the agreement of thermal limits operative in the field with the lab-determined pejus range supports previous studies, which interpreted thermal limitation in aquatic ectotherms to start with limited oxygen supply capacity (*3*–*6*, *16*–*18*) (Fig. 1, C to E). Accordingly, pejus limits are the earliest limits experienced by the whole organism in the field. Decrements in aerobic performance cause reduced growth and enhanced mortality first among larger specimens. A reduction in abundance results when all size groups of a population are affected. Residual variability in the data suggests that not only the temperature value itself but also the length of exposure is crucial in setting mortality. The data did not reveal an influence of the shift to milder winters. Further-

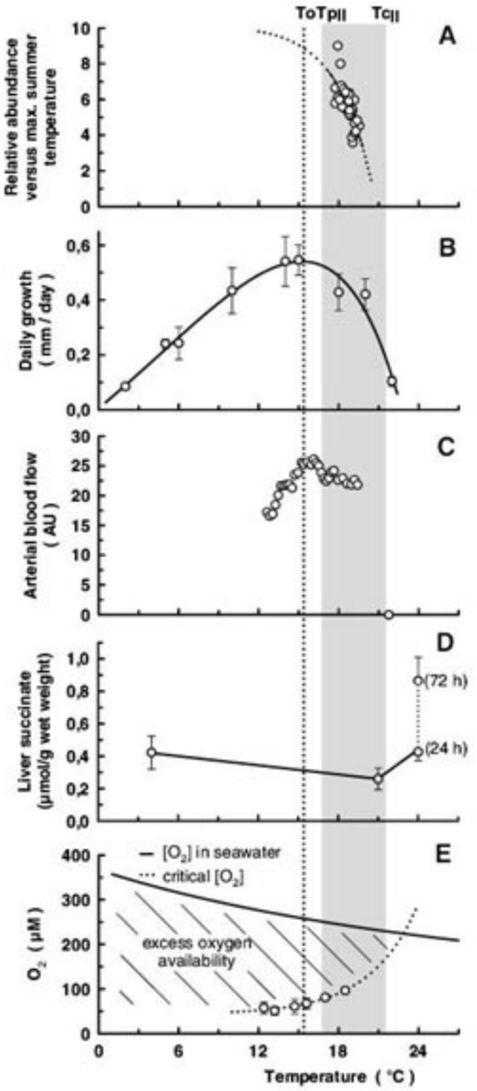


Fig. 1. Matching field and laboratory data reflect thermal limitation in eelpout in accordance with fig. S1. The shaded area characterizes the pejus range between upper T_p and T_c. (**A**) The negative correlation between summer water temperatures and relative abundance indicates heat-induced mortality of eelpout in the Wadden Sea [5-year running mean, recalculated from (*14*, *15*) and weather data licensed by DWD (Deutscher Wetterdienst)]. Data were fitted to $Ar(T) = Ar_{max}[1 - e^{k(T-T_o)}]$ where $Ar(T)$ is the relative abundance depending on temperature ($Ar_{max} = 10.173$; $k = 0.377$; $T_o = 20.853$, $r = 0.7130$, and $P < 0.01$). (**B**) Daily growth increments recalculated from (*31*) in relation to water temperature (mean ± SD). Data were fitted to the equation $dL(T) = F_1(T) + F_2(T) = (A_1 e^{B_1 T} + C_1) + (A_2 e^{B_2 T} + C_2)$ with $dL(T)$ = daily growth rates at maximum food supply. The first term, $F_1(T) = A_1 e^{B_1 T} + C_1$, represents the temperature dependence of aerobic processes supporting growth performance. The second term, $F_2(T) = A_2 e^{B_2 T} + C_2$, represents the parallel, exponential rise in processes limiting aerobic scope and thus growth capacity ($A_1 = 0.9901$, $B_1 = 0.0667$, $C_1 = -0.3953$, $A_2 = -0.1942$, $B_2 = 0.1299$, $C_2 = -0.3953$, $r = 0.9823$, and $P < 0.01$). (**C**) Arterial blood flow in relative units [measured by nuclear magnetic resonance (NMR) imaging techniques and recalculated as running means from (*4*)] reflects maximized circulatory oxygen supply at optimum temperature T_o and capacity limitation beyond T_{pII}. AU, arbitrary units. (**D**) Mismatch between oxygen supply and demand finally leads to the accumulation of succinate in the liver beyond T_{cII} [data are mean ± SD from (*32*)]. (**E**) At upper T_c, $[O_2]_c$ of the eelpout reaches air saturation oxygen concentrations. Water O_2 concentration (solid line) and $[O_2]_c$ (open circles, mean ± SE, fitted by dotted line) were recalculated from (*25*) for a salinity of 32 ‰. $[O_2]_c$ was fitted to $[O_2]_c(T) = C_1 T^{C2} + C_3$ ($C_1 = 4.33 \times 10^{-6}$, $C_2 = 5.56$, $C_3 = 48.19$, $r = 0.9764$, and $P < 0.01$). Reported data are valid for the largest specimens [between 23- and 25-cm body lengths (*25*)] found in the Wadden Sea.

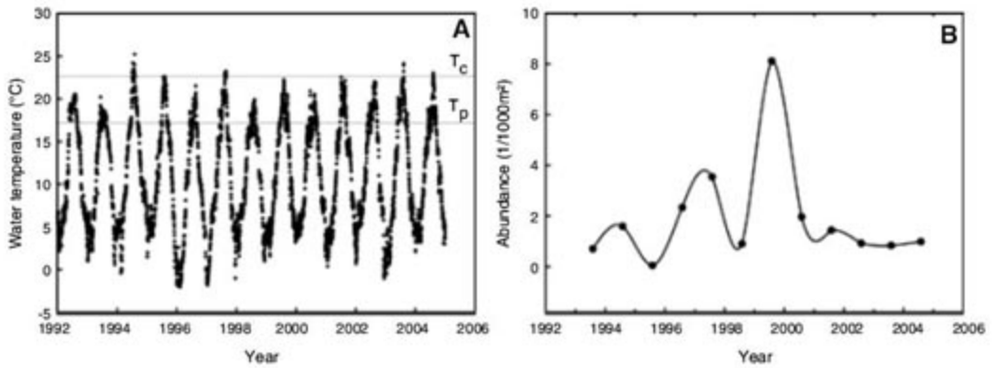


Fig. 2. (**A**) Daily water temperature (°C) in the Wadden Sea from 1992 to 2004 (DWD, Station Norderney). Putative upper critical (T_c) and pejus (T_p) temperatures of large eelpout (Fig. 1) are indicated by horizontal lines. (**B**) Total abundance (1/1000 m^2) of eelpout in the Wadden Sea sampling area during summers between 1993 and 2005. Data fitted by spline curve [Sigmaplot (*33*)].

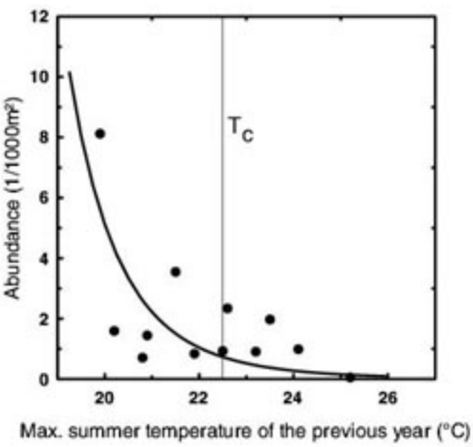


Fig. 3. Abundance of eelpout versus maximum summer temperature from the previous year (data from Fig. 2). Data were fitted to $A(T_{max-1}) = C_1 e^{C_2/C_3 T}$ where $A(T_{max-1})$ is the abundance depending on maximum temperature of the previous summer (fitting parameters $C_1 = 1.2 \times 10^{-7}$, $C_2 = 240.83$, $C_3 = 0.6686$, $r = 0.6599$, and $P < 0.01$). The putative upper T_c (Fig. 1) is indicated by a vertical line.

more, population growth depends on food supply, which in turn influences aerobic performance and thermal sensitivity. Potential additional components in field tolerance still need to be identified. However, we suggest that reduced aerobic performance beyond pejus limits enhances sensitivity to other, more obvious mechanisms eliciting mortality (predation, starvation, or disease). These influences would display their inherent variability and thereby enhance the variability in the temperature-dependence of abundance.

Matching thresholds in field and laboratory data highlight the ecological relevance of the concept of oxygen- and capacity-limited thermal tolerance. Adaptation to climate variability involves adjustments of functional capacity in general and, specifically, in the components of aerobic metabolism, of oxygen supply capacity, and of associated costs. Trade-offs in thermal adaptation processes and in organismic energy budget shape the width of thermal windows, with consequences for biogeography (*1*), growth performance, development, fecundity, recruitment, life-styles, and life-history evolution (*9*, *10*, *26*, *27*). Widths and locations of tolerance windows on the temperature scale may change or shift during ontogeny (*26*). At the ecosystem level, species-specific biogeographical ranges differ but overlap and imply that variable thermal windows and sensitivities cause variability in distribution shifts (*1*), species composition, seasonal timing, and associated mismatch phenomena in species interactions as in a food web. For example, the shift from larger (*Calanus finmarchicus*) to smaller (*C. helgolandicus*) copepod fauna in the southern North Sea caused reduced food availability for Atlantic cod (*G. morhua*) (*28*). This regime shift was largely determined by different thermal windows of the two copepod species (*29*). Warming-induced reductions of cod abundance are thus caused both directly [via thermal sensitivity of cod (*30*)] and indirectly [via the food web (*28*)] but based on the same physiological principles. Overall, the concept of oxygen- and capacity-limited thermal tolerance can provide an integrative framework for developing a cause-and-effect understanding of the influence of climate change and variability on marine ecosystems, including food web structure, recruitment success, and fish landings (*30*).

References and Notes

1. A. L. Perry, P. J. Low, J. R. Ellis, J. D. Reynolds, *Science* **308**, 1912 (2005); published online 12 May 2005 (10.1126/science.1111322).
2. M. N. Jensen, *Science* **299**, 38 (2003).
3. H. O. Pörtner, *Naturwissenschaften* **88**, 137 (2001).
4. F. C. Mark *et al.*, *Am. J. Physiol.* **283**, R1254 (2002).
5. M. Frederich, H. O. Pörtner, *Am. J. Physiol.* **279**, R1531 (2000).
6. H. O. Pörtner, *Comp. Biochem. Physiol. A* **132**, 739 (2002).
7. I. A. Johnston *et al.*, *J. Exp. Biol.* **201**, 1 (1998).
8. I. Hardewig *et al.*, *J. Comp. Physiol. B* **169**, 597 (1999).
9. H. O. Pörtner *et al.*, *Fish Physiol.* **22**, 79 (2005).
10. H. O. Pörtner, *Deep-Sea Res. II* **53**, 1071 (2006).
11. K. H. Wiltshire, B. F. J. Manly, *Helgol. Mar. Res.* **58**, 269 (2004).
12. C. Sheppard, *Mar. Pollut. Bull.* **49**, 12 (2004).
13. Z. W. Kundzewicz, M. L. Parry, in *IPCC: Climate Change 2002: Impacts, Adaptation, and Vulnerability*, J. J. McCarthy *et al.*, Eds. (Cambridge Univ. Press, Cambridge, 2001), pp. 641–692.
14. K. Tiews, *Arch. Fischereiwiss.* **34**, 4 (1983).
15. K. Tiews, H. Wienberg, *Veröff. Inst. Küsten Binnenfisch. Hamburg* **103**, 1 (1990).
16. H. O. Pörtner *et al.*, *Respir. Physiol. Neurobiol.* **141**, 243 (2004).
17. G. Lannig *et al.*, *Am. J. Physiol.* **287**, R902 (2004).
18. F. J. Sartoris *et al.*, *J. Fish Biol.* **62**, 1239 (2003).
19. Materials and methods are available on *Science* Online.
20. J. Ulleweit, thesis, University at Bremen, Bremen, Germany (1995).
21. S. M. Sogard, *Bull. Mar. Sci.* **60**, 1129 (1997).
22. H. O. Pörtner, *Physiol. Biochem. Zool.* **77**, 959 (2004).
23. T. K. Linton *et al.*, *Can. J. Fish. Aquat. Sci.* **55**, 576 (1998).
24. K. J. Rodnick *et al.*, *J. Fish Biol.* **64**, 310 (2004).
25. M. V. Zakhartsev *et al.*, *J. Comp. Physiol. B* **173**, 365 (2003).
26. H. O. Pörtner *et al.*, *Physiol. Biochem. Zool.* **79**, 295 (2006).
27. H. O. Pörtner, *J. Exp. Biol.* **205**, 2217 (2002).
28. G. Beaugrand *et al.*, *Nature* **426**, 661 (2003).
29. P. Helaouët, G. Beaugrand, *Mar. Ecol. Progr. Ser.*, in press.
30. H. O. Pörtner *et al.*, *Cont. Shelf Res.* **21**, 1975 (2001).
31. M. Fonds *et al.*, (Demersal Fish Committee), *Metabolism, Food Consumption, Growth, and Food Conversion of Shorthorn Sculpin* (Myoxocephalus scorpius) *and Eelpout* (Zoarces viviparus) (Document G:31, International Council for Exploration of the Sea, Copenhagen, 1989).
32. P. L. M. van Dijk *et al.*, *J. Exp. Biol.* **202**, 3611 (1999).
33. Sigmaplot, Systat Software Incorporated, San Jose, CA.

34. Supported by the MarCoPoli program of the Alfred Wegener Institute.

Supporting Online Material

www.sciencemag.org/cgi/content/full/315/5808/95/DC1
Materials and Methods
Figs. S1 and S2

22 September 2006; accepted 29 November 2006
10.1126/science.1135471

A Hexanucleotide Element Directs MicroRNA Nuclear Import

Hun-Way Hwang,[1] Erik A. Wentzel,[2] Joshua T. Mendell[1,2]*

MicroRNAs (miRNAs) negatively regulate partially complementary target messenger RNAs. Target selection in animals is dictated primarily by sequences at the miRNA 5′ end. We demonstrated that despite their small size, specific miRNAs contain additional sequence elements that control their posttranscriptional behavior, including their subcellular localization. We showed that human miR-29b, in contrast to other studied animal miRNAs, is predominantly localized to the nucleus. The distinctive hexanucleotide terminal motif of miR-29b acts as a transferable nuclear localization element that directs nuclear enrichment of miRNAs or small interfering RNAs to which it is attached. Our results indicate that miRNAs sharing common 5′ sequences, considered to be largely redundant, might have distinct functions because of the influence of cis-acting regulatory motifs.

Nucleotides 2 to 7 of microRNAs (miRNAs), known as "seed" sequences, are considered the most critical for selecting targets. Within a given species, highly related miRNAs sharing a common seed sequence are grouped into miRNA families, are predicted to have overlapping targets, and are considered to be largely redundant (*1*–*5*). Nevertheless, loss of function of miRNA family members with divergent 3′ end sequences results in overlapping but distinct phenotypes in *Caenorhabditis elegans* and in *Drosophila* (*6*, *7*). These distinct phenotypes often do not appear to be due to differences in miRNA expression patterns, which raises the possibility that distinct sequences within miRNA family members confer upon them characteristic functional properties. Here we describe a sequence motif that dramatically influences the posttranscriptional behavior of a human miRNA.

Examination of cell-cycle stage–specific miRNA expression patterns with a previously described oligonucleotide array (*8*) revealed substantial accumulation of miR-29 in mitotic HeLa cells (*9*). There are three human miR-29 paralogs: miR-29a, miR-29b, and miR-29c (fig. S1A). A highly specific Northern blot assay (fig. S1B) demonstrated that each exhibits a distinct expression pattern. miR-29a is constitutively expressed in all cell-cycle phases, miR-29b is present at low levels except in mitotic cells, and miR-29c is not detectable (Fig. 1A).

Human miR-29 family members are encoded by the miR-29b-1/miR-29a cluster and the miR-29b-2/miR-29c cluster (Fig. 1B). A fragment encompassing the miR-29b-1/miR-29a cluster was amplified by reverse transcription polymerase chain reaction (RT-PCR) after small interfering RNA (siRNA)–mediated inhibition of Drosha (which performs the first step in

[1]Program in Human Genetics and Molecular Biology, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. [2]McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA.

*To whom correspondence should be addressed. E-mail: jmendell@jhmi.edu

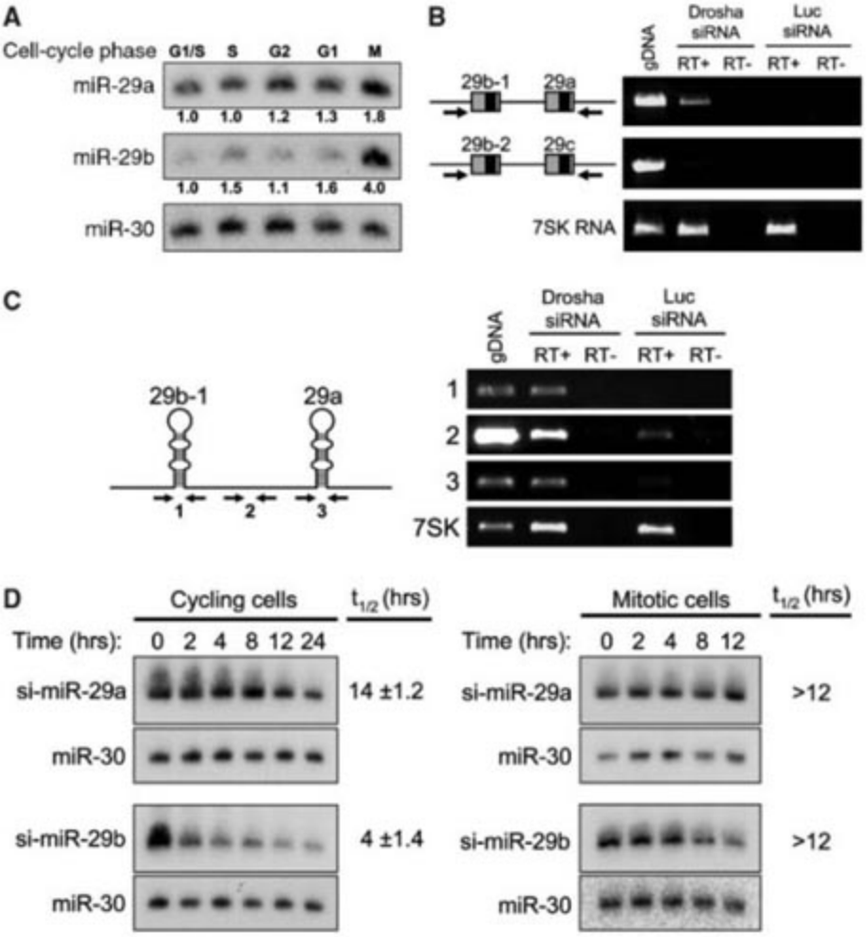


Fig. 1. miR-29b is degraded rapidly in cycling cells but is stable in mitotic cells. (**A**) Northern blot demonstrating discordant expression of miR-29a and miR-29b during the cell cycle. miR-29c (not shown) was not detectable. miR-30, a constitutively expressed miRNA, served as a loading control. Relative expression levels are indicated below the two panels. (**B**) The miR-29b-1/miR-29a cluster but not the miR-29b-2/miR-29c cluster is expressed in HeLa cells. The genomic organization of each cluster is shown on the left (arrows represent primers). Genomic DNA (gDNA) or RNA from cells treated with siRNA directed against Drosha or Luc was amplified. 7SK RNA served as a positive control. (**C**) RT-PCR assay for Drosha processing. Numbers in the schematic indicate amplicons. Primer pairs, indicated by arrows in the schematic and by numbers to the left of the gels, were used to amplify gDNA or RNA from cells treated with siRNA directed against Drosha or Luc. (**D**) Pulse-chase assay to measure synthetic miRNA stability. Northern blotting was used to measure the abundance of si–miR-29a and si–miR-29b at the indicated time points after transfection of RNA duplexes. The mean and SD of calculated half-lives ($t_{1/2}$) from three experiments are shown on the right of each blot. The mitotic half-lives were longer than the time course and are therefore reported as >12 hours.

miRNA processing), demonstrating that these miRNAs are cotranscribed as a polycistronic primary transcript. In contrast, neither the miR-29b-2/miR-29c cluster primary transcript nor mature miR-29c was detected by RT-PCR or Northern blotting. These results indicate that miR-29b likely derives exclusively from the miR-29b-1/miR-29a cluster in HeLa cells. Moreover, given that miR-29a and miR-29b are cotranscribed and miR-29a is constitutively expressed, a posttranscriptional mechanism must be functioning to prevent the accumulation of miR-29b in all cell-cycle phases except mitosis.

Posttranscriptional regulation of miRNA abundance could occur at the level of miRNA maturation or stability. miRNA maturation is a two-step process involving sequential cleavages by Drosha and Dicer (*10*, *11*). To assess Drosha processing of miR-29b-1, we performed RT-PCR with an amplicon spanning the pre-miRNA hairpin (Fig. 1C, amplicon 1). As expected, an RT-PCR product was observed only after siRNA-mediated depletion of Drosha. Failure to amplify this region in the presence of Drosha activity was not due to a general instability of the primary transcript because an amplicon that did not span a Drosha cleavage site (Fig. 1C, amplicon 2) produced a product with control [luciferase (Luc)] or Drosha siRNA treatment. Two lines of evidence also suggest efficient processing of miR-29b by Dicer. First, we never observed accumulation of pre–miR-29b, a species that frequently accumulates when Dicer processing is blocked (*12*). Second, a fully processed synthetic siRNA-like miR-29b duplex (si–miR-29b) mimicked the endogenously expressed miRNA and accumulated in cells arrested in mitosis (fig. S2). Thus, the mitotic accumulation of miR-29b occurs after it is fully processed and therefore is most likely due to enhanced stability of the miRNA in this cell-cycle phase.

In order to measure miRNA stability in cycling and mitotic cells, we used a pulse-chase strategy with synthetic miRNA duplexes

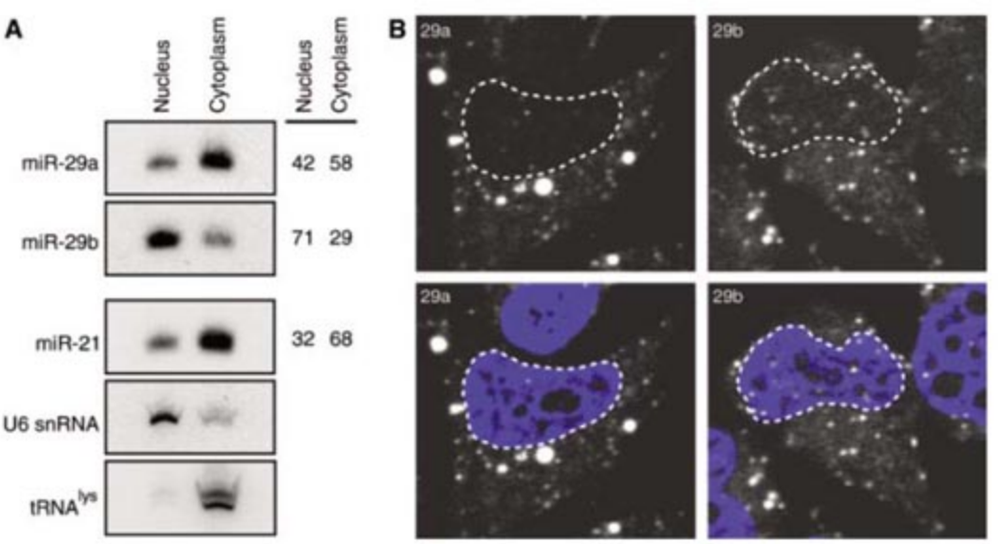


Fig. 2. miR-29b is imported into the nucleus. (**A**) Northern blot demonstrating the relative nuclear and cytoplasmic abundance of endogenous miR-29a, miR-29b, and miR-21. U6 small nuclear RNA (U6 snRNA) and lysine-tRNA ($tRNA^{lys}$) served as controls for subcellular fractionation. (**B**) Confocal microscopy demonstrating intranuclear localization of a fraction of si–miR-29b but not si–miR-29a (upper panels). The dashed line indicates the nuclear periphery, as defined by 4′,6′-diamidino-2-phenylindole staining (lower panels).

(si-miRNAs). Multiple lines of evidence indicate that si–miR-29b accurately recapitulates the behavior of the endogenous miRNA. First, as mentioned above, si–miR-29b exhibits mitotic accumulation like the endogenous miRNA (fig. S2). Second, immunoprecipitation experiments with human Argonaute 1 or 2 demonstrate that si–miR-29b duplexes are appropriately loaded into the RNA-induced silencing complex (fig. S3). Cells were pulsed with si-miRNAs, and the fraction that remained was monitored over time (Fig. 1D). si–miR-29b degraded rapidly in cycling cells but was stable in mitotic cells. As expected, si–miR-29a degraded slowly under both tested conditions.

In these studies, nocodazole treatment was used to obtain cells arrested in mitosis. Several experiments were performed to demonstrate that nocodazole-induced accumulation of miR-29b was a result of mitotic arrest rather than nonspecific perturbation of microtubule dynamics (supporting online material (SOM) text and fig. S4). For example, miR-29b accumulated only when cells entered mitosis despite being exposed to nocodazole throughout a prolonged time course (fig. S4A). Additionally, we demonstrated that the accumulation of miR-29b was not related to apoptosis, which is known to occur when cells are arrested at the spindle checkpoint (fig. S4D) (*13*).

miR-29a and all previously studied animal miRNAs are predominantly cytoplasmic (Fig. 2A) (*14*, *15*). Because disassembly of the nuclear membrane distinguishes mitosis from other cell-cycle phases, we examined the subcellular localization of miR-29b. Cellular fractionation revealed that miR-29b (which exists at low but detectable levels in cycling cells) is predominantly nuclear. Synthetic si–miR-29a and si–miR-29b exhibited the identical localization pattern as their native counterparts (Fig. 3). A plasmid expressing a fragment of the miR-29b-1 primary transcript also produced mature miR-29b that trafficked to the nucleus (fig. S5). Fluorescence in situ hybridization and confocal microscopy demonstrated punctate cytoplasmic localization of transfected si–miR-29a or si–miR-29b duplexes in a pattern that is consistent with other published reports (Fig. 2B) (*16*). In addition, substantially more punctate and diffuse intranuclear staining was visible for si–miR-29b. These data indicate that miR-29b is imported into the nucleus in cycling cells. miR-29b also shows mitotic accumulation and nuclear enrichment in murine NIH 3T3 cells (fig. S6), demonstrating the conservation of this pathway in other mammalian cell lines.

Because si–miR-29a and si–miR-29b behave exactly like the endogenously expressed miRNAs, sequence elements contained within the fully processed molecules must specify their distinct localization patterns. A uridine at nucleotide 10 and a distinctive hexanucleotide 3′ terminal motif (AGUGUU) distinguish miR-29b from miR-29a (Fig. 3). Variant si–miR-29a duplexes containing a uridine at position 10 [si–miR-29a(C10U)] were cytoplasmic, whereas si–miR-29a duplexes tagged with the miR-29b 3′ terminal motif [si–miR-29a(TM^{29b})] were enriched in the nucleus (Fig. 3). To determine whether this motif could confer nuclear localization to an unrelated sequence, we assessed its ability to direct nuclear import of a functional siRNA directed against luciferase (*17*). Addition of the miR-29b terminal motif was sufficient to direct nuclear enrichment [Fig. 3, si-Luc(TM^{29b})]. An additional unrelated siRNA ending in AGUGUU was tested and found to be highly enriched in the nuclear compartment [Fig. 3, si-test(TM^{29b})]. Internal deletions introduced into si–miR-29b demonstrated that the motif can specify nuclear import of 22- or 21-nucleotide miRNAs (Fig. 3). However, the motif must be at the 3′ terminus of the miRNA in order to function [fig. S7, si–miR-29b(TM^{-1}) and si–miR-29b(TM^{-2})]. Finally, we determined the consequences of transversion mutations at each position of the element. Mutations at four positions substantially reduced the nuclear-targeting efficiency (fig. S7). The miR-29b terminal motif or a relaxed consensus supported by these mutagenesis studies (AGNGUN, where N is any nucleotide) is not present in other miRNAs that are conserved throughout the mammalian radiation (SOM text). Therefore, the mir-29b terminal motif appears to be rarely, if ever, used by other mammalian miRNAs for nuclear localization.

Although miR-29b undergoes rapid decay, accelerated turnover does not appear to be a general feature of small RNAs that are imported into the nucleus. The addition of the miR-29b terminal motif did not lead to the accelerated decay of miR-29a or Luc siRNA (fig. S8). The introduction of the C-to-U substitution at position 10 into miR-29a [si–miR-29a(C10U)] resulted in a shorter half-life but did not fully recapitulate the rapid decay observed for miR-29b (fig. S8). These findings indicate that the miR-29b terminal motif may be useful for designing stable siRNAs or miRNAs that are highly enriched in the nucleus. These RNAs may prove useful for the manipulation of nuclear steps in gene expression. Indeed, it is possible that the natural function of miR-29b is to regulate transcription or splicing of target transcripts, rather than the canonical translation regulatory functions that are ascribed to other miRNAs.

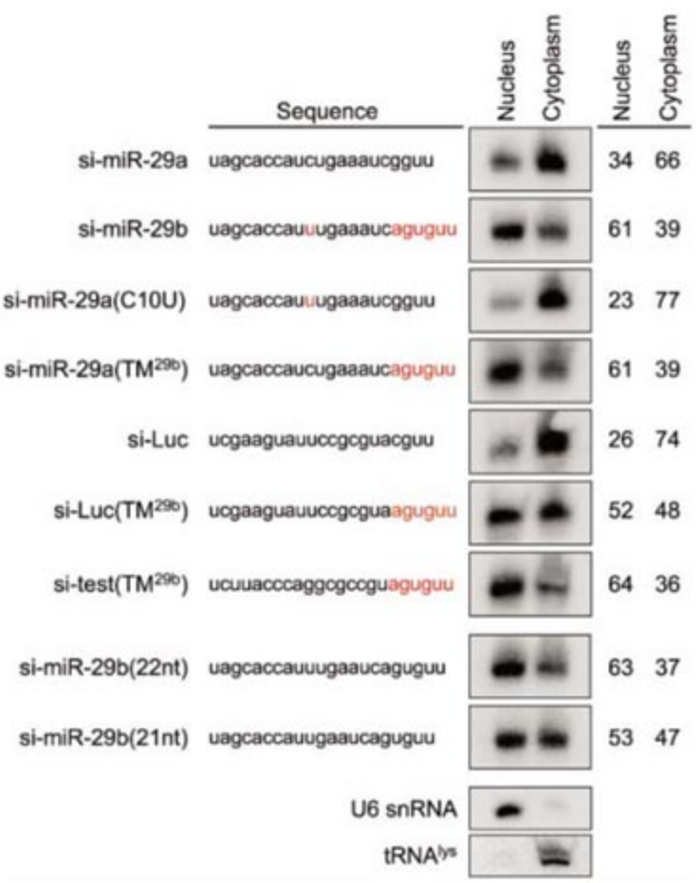


Fig. 3. The miR-29b 3′ terminal motif is a transferable nuclear import element. Northern blotting was used to determine the nuclear and cytoplasmic abundance of a series of synthetic siRNAs. All membranes were reprobed with U6 snRNA and lysine-tRNA to verify successful subcellular fractionation (representative blots shown). nt, nucleotide. Red letters in sequences indicate substitutions or altered positions.

References and Notes

1. M. A. Valencia-Sanchez, J. Liu, G. J. Hannon, R. Parker, *Genes Dev.* **20**, 515 (2006).
2. T. Du, P. D. Zamore, *Development* **132**, 4645 (2005).
3. J. Brennecke, A. Stark, R. B. Russell, S. M. Cohen, *PLoS Biol.* **3**, e85 (2005).
4. E. C. Lai, *Genome Biol.* **5**, 115 (2004).
5. B. P. Lewis, C. B. Burge, D. P. Bartel, *Cell* **120**, 15 (2005).
6. D. Leaman *et al.*, *Cell* **121**, 1097 (2005).
7. A. L. Abbott *et al.*, *Dev. Cell* **9**, 403 (2005).
8. K. A. O'Donnell, E. A. Wentzel, K. I. Zeller, C. V. Dang, J. T. Mendell, *Nature* **435**, 839 (2005).
9. Materials and methods are available as supporting material on *Science* Online.
10. B. R. Cullen, *Mol. Cell* **16**, 861 (2004).
11. V. N. Kim, *Nat. Rev. Mol. Cell. Biol.* **6**, 376 (2005).
12. J. M. Cummins *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 3687 (2006).
13. M. Castedo *et al.*, *Oncogene* **23**, 2825 (2004).

14. G. Meister *et al.*, *Mol. Cell* **15**, 185 (2004).
15. Y. Lee, K. Jeon, J. T. Lee, S. Kim, V. N. Kim, *EMBO J.* **21**, 4663 (2002).
16. K. M. Pauley *et al.*, *EMBO Rep.* **7**, 904 (2006).
17. S. M. Elbashir *et al.*, *Nature* **411**, 494 (2001).
18. The authors thank T. Tuschl for plasmids; M. Delannoy for assistance with microscopy; F. Spencer for helpful discussions; and M. Awad, J. Boeke, H. Dietz, R. Green, K. O'Donnell, and members of the Mendell Lab for critical reading of the manuscript. J.T.M. is a March of Dimes Basil O'Connor Scholar and a Rita Allen Foundation Scholar and receives support from the Lustgarten Foundation for Pancreatic Cancer Research.

Supporting Online Material

www.sciencemag.org/cgi/content/full/315/5808/97/DC1
Materials and Methods
SOM Text
Figs. S1 to S9
Table S1 to S5
References

12 October 2006; accepted 17 November 2006
10.1126/science.1136235

Left-Right Dynein Motor Implicated in Selective Chromatid Segregation in Mouse Cells

Athanasios Armakolas* and Amar J. S. Klar†

During cell division, copies of mouse chromosome 7 are segregated selectively or randomly to daughter cells depending on the cell type. The mechanism for differential segregation is unknown. Because mouse left-right dynein (LRD) gene mutations result in randomization of visceral organs' laterality, we hypothesized that LRD may also function in selective chromatid segregation. Indeed, upon knock-down by RNA interference methods, LRD depletion disrupts biased segregation. LRD messenger RNA presence or absence correlates with the observed segregation patterns. This work supports the claim that LRD functions in a mechanism for selective chromatid segregation.

Products of DNA by semiconservative replication (*1*) include one chromatid possessing an "older Watson" (W) strand (*2*) and a newly synthesized complementary Crick strand (C′), whereas the sister chromatid contains an "older Crick" (C) strand and a newly synthesized complementary Watson strand (W′) (fig. S1). The WW:CC designation is used when both older W-containing chromatids, copied from a homologous pair of chromosomes, are delivered to one daughter cell and both older C-containing chromatids are delivered to the other daughter cell in mitosis (fig. S1). The WC:WC designation is used in the segregation of an older W-containing plus an older C-containing chromatid pair to both daughter cells (*3*). In a recent study, we observed nonrandom segregation of chromatids of chromosome (Chr) 7 in mouse embryonic stem cells (ES), endoderm, and neuroectoderm cells supporting a mechanism of nonrandom segregation, whereas pancreatic, mesoderm, and cardiomyocyte cells displayed random segregation (*4*). The mechanism and physiological relevance for nonrandom chromatid segregation is currently unknown (*5*, *6*).

Mutations of the *inversus viscerum* (*iv*) locus of mice cause left-right (LR) axis randomization (*7*). We hypothesized that the *iv* locus dictates biased DNA strand/chromatid segregation (*8*). The *iv* locus encodes a microtubule-based motor protein; it is named left-right dynein (LRD) (*9*). Its molecular role in LR axis specification remains controversial (*10*). We examine here the effect of the *LRD* gene on nonrandom chromatid segregation.

Details of the site-specific recombination model we employed and its features allowing us to deduce the mode of specific chromatid distribution are presented in (*11*). The *LRD* gene mRNA was knocked down (KD) in cells using small-interfering RNA (siRNA) technology. The identification of which chromatid segregated to the *HAT*^{r} (hypoxanthine, adenine, and thymidine) daughter cell that resulted from experimentally induced interchromosomal mitotic recombination was made by determining which Chr 7 markers, *M* and *P*, were inherited (fig. S1). The markers' constitution was determined by Southern blot analysis (*11*). Results with control lines confirm results of earlier findings (*4*, *12*); all 20 ES and all 20 endoderm recombinants tested were *M/M* homozygous (Fig. 1 and table S1), reflecting the preferential WW:CC segregation mode (fig. S1), and all 20 neuroectoderm recombinants remained *M/P* heterozygous, possibly indicating the biased WC:WC mode (fig. S1). Interestingly, both biased modes were disrupted in the *LRD* KD lines (Fig. 1 and table S1).

Gene Regulation and Chromosome Biology Laboratory, Center for Cancer Research, National Cancer Institute at Frederick, Post Office Box B, Frederick, MD 21702–1201, USA.

*Present address: Hippokrateion Hospital of Athens, Laboratory of Surgical Research, Vassilisis Sofias 114, Athens 11527, Greece.

†To whom correspondence should be addressed. E-mail: klar@ncifcrf.gov

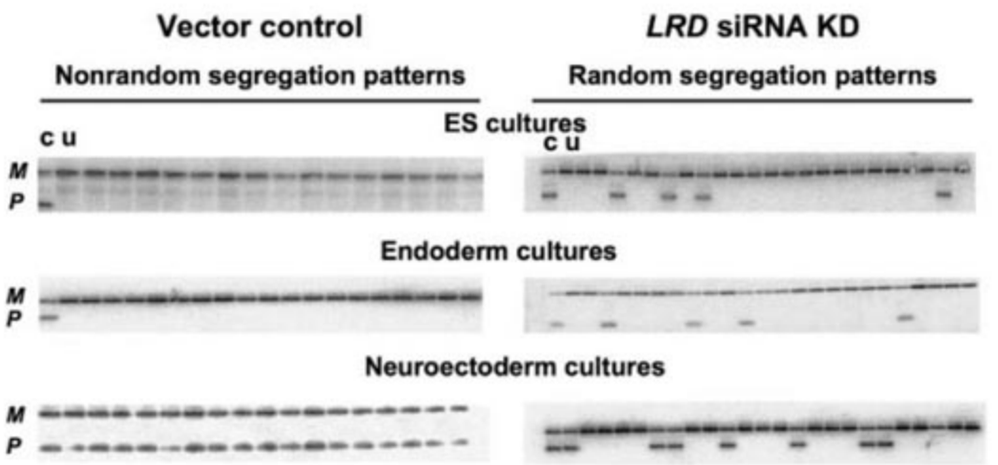


Fig. 1. Southern blot analysis indicates Chr 7 alleles composition [see (*11*) for details]. Lane c represents a control culture showing signals for *M* (maternal chromosome) and *P* (paternal chromosome) allelic markers. Lane u (uninfected) control represents a *HAT*^{r} recombinant culture of the line that was not infected with the virus. Remaining lanes represent individual recombinant cultures that were selected from a line infected with the indicated virus. Analysis of the first 23 recombinants from a total of 50 examined (table S1) from *LRD* KD lines is displayed on the right.

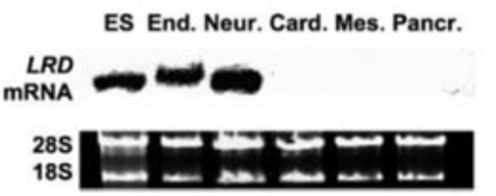


Fig. 2. Northern blot analysis of *LRD* mRNA [see (*11*) for details]. Ethidium bromide staining of 28S and 18S ribosomal RNA species serves as RNA loading control. ES, embryonic stem cells; End., endoderm; Neur., neuroectoderm; Card., cardiomyocytes; Mes., mesoderm; Pancr., pancreatic cells.

The mechanism for segregation specificity by cell type could stem from regulation of mRNA expression or stability. *LRD* mRNA was detected (Fig. 2) in ES, endoderm, and neuroectoderm cultures, which follow the biased pattern, but not in pancreatic, mesoderm, and cardiomyocyte cultures, which follow the random pattern (*4*).

Our results confirm the existence of the biased chromatid segregation phenomenon (*3*, *4*, *8*) and provide a mechanism for how the cell type regulates the segregation specificity. The *LRD* KD disrupts two types of nonrandom segregation modes in three lineages tested, and *LRD* mRNA abundance correlates with the segregation mode in each of six cell lineages examined. Considering these results, LRD likely functions in the nonrandom segregation mechanism.

References and Notes

1. M. Meselson, F. W. Stahl, *Proc. Natl. Acad. Sci. U.S.A.* **44**, 671 (1958).
2. J. D. Watson, F. H. Crick, *Nature* **171**, 737 (1953).
3. A. J. S. Klar, *Genetics* **167**, 1833 (2004).
4. A. Armakolas, A. J. S. Klar, *Science* **311**, 1146 (2006).
5. J. E. Haber, *Science* **313**, 1045b (2006).
6. A. J. S. Klar, A. Armakolas, *Science* **313**, 1045c (2006).
7. K. P. Hummel, D. B. Chapman, *J. Hered.* **50**, 9 (1959).
8. A. J. Klar, *Trends Genet.* **10**, 392 (1994).
9. D. M. Supp, D. P. Witte, S. S. Potter, M. Brueckner, *Nature* **389**, 963 (1997).
10. C. Tabin, *J. Mol. Histol.* **36**, 317 (2005).
11. Materials and methods are available as supporting material on *Science* Online.
12. P. Liu, N. A. Jenkins, N. G. Copeland, *Nat. Genet.* **30**, 66 (2002).

13. We thank M. Lewandoski [National Cancer Institute (NCI)] for the tissue culture facility use and S. Austin (NCI) for improving the presentation of the paper. Editorial assistance of the NCI Fellows Editorial Board is appreciated. The Intramural Research Program of NCI supported this research.

Supporting Online Material

www.sciencemag.org/cgi/content/full/315/5808/100/DC1
Materials and Methods
SOM Text
Figs. S1 and S2
Table S1
References

2 May 2006; accepted 10 November 2006
10.1126/science.1129429

A Cytokinin Perception Mutant Colonized by *Rhizobium* in the Absence of Nodule Organogenesis

Jeremy D. Murray,[1] Bogumil J. Karas,[1] Shusei Sato,[2] Satoshi Tabata,[2] Lisa Amyot,[1] Krzysztof Szczyglowski[1]*

In legumes, Nod-factor signaling by rhizobia initiates the development of the nitrogen-fixing nodule symbiosis, but the direct cell division stimulus that brings about nodule primordia inception in the root cortex remains obscure. We showed that *Lotus japonicus* plants homozygous for a mutation in the *HYPERINFECTED 1* (*HIT1*) locus exhibit abundant infection-thread formation but fail to initiate timely cortical cell divisions in response to rhizobial signaling. We demonstrated that the corresponding gene encodes a cytokinin receptor that is required for the activation of the nodule inception regulator *Nin* and nodule organogenesis.

The development of nitrogen-fixing nodules on roots of leguminous plants commences with a molecular dialogue between the host plant and a compatible strain of rhizobia, leading to the synthesis of bacterially encoded lipochito-oligosaccharide signaling molecules, the Nod factors (NFs). Plant plasma membrane–derived structures called infection threads (ITs), which originate within root hairs of the host root in a NF-dependent manner, act as conduits for rhizobia to enter the root tissues and to progress toward the root cortex where a nodule primordium (NP) has been initiated. The release of bacteria from ITs into the cytosol of a subset of NP cells and subsequent cellular specialization of both symbionts lead to the formation of fully functional nitrogen-fixing organs, the root nodules (*1*).

The initiation of cell divisions for NP organogenesis is presumed to arise from the relay of a signal from the epidermally perceived NFs to the root cortex. Several genes in the NF-dependent signaling pathway have been identified, including putative NF receptors (*2*–*4*) and a number of downstream elements (*5*–*12*). Deleterious mutations in any of these genes abolish bacterial entry into the root and NP development, indicating a crucial role for NF signaling in both processes.

The identification of plants that spontaneously form nodules (*13*–*15*), together with the observations that ectopic application of cytokinins (*16*) or auxin transport inhibitors (*17*) to the root surface lead to the development of nodule-like structures, demonstrates that the machinery required for NP development is intrinsic to the plant. Consequently, the NF pathway is presumed to trigger nodule organogenesis by regulating the endogenous plant mechanism; however, the nature of the mechanism involved remains unclear.

We performed a screen for genetic suppressors of the *Lotus japonicus har1-1* hypernodulation phenotype that identified three allelic suppressor lines characterized by a low-nodulation phenotype and an excessive formation of ITs. The corresponding locus was named *HYPERINFECTED 1* (*HIT1*) (*18*). Further detailed phenotypic analysis, performed in both double (*hit1 har1-1*) and single (*hit1*) mutant backgrounds, showed that the three suppressor lines had indistinguishable mutant phenotypes, with the bacterial root invasion by way of ITs intact and the timely onset of associated cortical cell divisions for NP organogenesis aborted. The *hit1-1 har1-1* and *hit1-1* mutants were chosen as reference lines.

When analyzed 10 days after inoculation (dai) with a *Mesorhizobium loti* strain carrying a constitutive *hemA::lacZ* reporter gene fusion, the most noticeable feature of *hit1-1 har1-1* roots was hyperinfection (Fig. 1A and fig. S1A). The large number of ITs that formed in the *hit1-1 har1-1* mutant roots originated within curled root hairs, but their progression toward the root cortex was blocked at the interface between the epidermis and the cortex (Fig. 1B). Infrequent ITs that escaped this early blockage and managed to penetrate within the *hit1-1 har1-1* root cortex looped frequently, suggesting that they were misguided (Fig. 1C). In spite of abundant infection events at the root epidermis, the root cortex of the *hit1-1 har1-1* mutant failed to initiate NP (Fig. 1C). By 14 dai, many ITs overcame the initial blockage and progressed deeper into the mutant root (Fig. 1D). Cortical cell divisions were initiated coincident with the accumulation of ITs within the root cortex, but NP did not develop (Fig. 1, D and E).

hit1-1 displayed the same mutant phenotypic features as *hit1-1 har1-1*, although the overall number of symbiotic events observed was reduced, likely reflecting the presence of the functional HAR1 autoregulatory receptor kinase (*19*). In *hit1-1*, an initial lack of NP formation in response to rhizobial infection (fig. S1B) was accompanied by the early onset of hyperinfection (fig. S1A) with a large number of ITs located within the root cortex (fig. S1C). Most root cortical cells associated with ITs in *hit1-1* and *hit1-1 har1-1* remained small and uncolonized; *M. loti* was confined to ITs (fig. S1D). Occasionally, a local release of bacteria from clustered ITs resulted in enlarged and often flattened

[1]Agriculture and Agri-Food Canada, Southern Crop Protection and Food Research Centre, 1391 Sandford Street, London, Ontario N5V 4T3, Canada. [2]Kazusa DNA Research Institute, Kisarazu, Chiba, 292-0812, Japan.

*To whom correspondence should be addressed. E-mail: szczyglowskik@agr.gc.ca

nodules (fig. S2, A and B), giving rise to the low-nodulation phenotype (fig. S2C). Intermittently, we also observed the development of nodules similar to those in the wild-type roots in both *hit1-1* and *hit1-1- har1-1* (fig. S2B).

We further investigated a role for the *HIT1* locus in NP organogenesis by studying *Early Nodulin 40* (*ENOD40*) and *Nin* expression, two markers for NP initiation (*12*, *20*). Quantitative reverse transcription polymerase chain reaction (qRT-PCR) analysis showed that steady-state levels of the corresponding transcripts in the inoculated roots were strongly reduced in the *hit1-1 har1-1* compared with those of the *har1-1* parental line, although this difference was not statistically significant between inoculated *hit1-1* and wild-type roots (fig. S3, A and B). We concluded that the presence of a functional *HIT1* locus is required for normal *ENOD40* and *Nin* expression during the organogenesis of NPs, at least in the *har1-1 hit1-1* background.

Additional validation of *HIT1* function in nodule organogenesis was provided by the analysis of the *L. japonicus root hairless* (*Ljrhl1-1*) and *hit1-1* double mutant. In response to inoculation with *M. loti*, *Ljrhl1-1* initially develops a large number of uncolonized NPs, a consequence of the absence of root hairs and associated ITs (*21*). We reasoned that if *HIT1* mediates nodule primordia organogenesis, the presence of a mutant *hit1-1* allele should prevent or notably attenuate NP formation in *Ljrhl1-1*. The analysis of the *Ljrhl1-1 hit1-1* mutant showed that the development of NP was almost entirely aborted, providing strong evidence for the indispensable role of *HIT1* in nodule organogenesis (Fig. 2). The *hit1-1* phenotype resembled the infection-thread "symbiosis" proposed to have been a precursor to nodulation (*22*). Could *HIT1* be the missing evolutionary link?

To begin addressing this question, we set out to map-base clone the *HIT1* locus. While this work was in progress, Jens Stougaard's group at Aarhus University (Denmark) cloned the *L. japonicus Spontaneous nodule formation 2* locus (*13*), renamed as *Lotus histidine kinase 1* (*Lhk1*) (*23*), which is localized to the same genetic interval on chromosome IV as *HIT1*. Given the opposing nodulation phenotypes of *snf2* and *hit1-1*, we tested the hypothesis that *hit1* mutants represent loss-of-function alleles of *Lhk1*.

Lhk1 specific primers were used to amplify the analogous genomic region from the three *hit1* lines. In all three lines, mutations were found that were predicted to result in premature stop codons (Fig. 3A). This finding, along with the ability of a wild-type *Lhk1* gene to complement the *hit1-1 har1-1* and *hit1-1* phenotypes in transgenic *L. japonicus* hairy root experiments (fig. S4), confirmed the identity of the underlying gene as *Lhk1*.

Because the *hit1-1* allele carried a G_{1695}-to-A nucleotide substitution in the splice donor site of intron four, we used oligonucleotide primers flanking this site to amplify the corresponding cDNA. Seven aberrant *hit1-1* splice variants were identified (Fig. 3B). In addition, a polymorphic species was found among the PCR products. A search for the corresponding *L. japonicus* genomic sequence identified a previously unrecognized gene, here named *Lhk2* (*Lotus histidine kinase 2*), of which the predicted product showed 85% identity at the amino acid level with LHK1. Whereas *Lhk1* transcripts were present in roots, nodules, and shoots (Fig. 3C), the *Lhk2* mRNA was detectable only in roots (fig. S5).

Analysis of the full-length *Lhk1* cDNA revealed a 2979–base pair open reading frame encoding a predicted protein of 993 amino acids (Fig. 3A). The LHK1 protein had 64% identity with the *Arabidopsis* cytokinin histidine kinase receptor AHK4, and 49 and 45% identity with AHK2 and AHK3, respectively. Similar to LHK1, LHK2 was more closely related to AHK4 than to other *Arabidopsis* cytokinin receptors (fig. S6).

When expressed in the *sln1*Δ yeast strain carrying a lethal mutation in SLN1 histidine kinase (*24*), *Lhk1* rescued the growth of the yeast strain in a cytokinin-dependent manner, demonstrating that LHK1 is a cytokinin receptor (Fig. 3D). In agreement with this notion, roots of *hit1-1* mutants exhibited strong insensitivity to exogenously applied cytokinin (Fig. 4A). A similar cytokinin-insensitive root phenotype was observed in all three *hit1-1 har1-1* double-mutant lines (fig. S7).

Because the accumulation of *Nin* and *ENOD40* transcripts was significantly attenuated in the *hit1-1 har1-1* mutant, we next tested whether exogenous application of cytokinin regulates expression of these genes in the wild-type roots. *ENOD40* has been shown to be induced by external application of cytokinin to the roots in several legume species (*25*), and this was also the case in *L. japonicus*, although the overall induction of *ENOD40* was rather modest (fig. S8A). In contrast, 50 nM benzyl adenine (BA) increased the steady-state level of *Nin* transcript by a factor of 20 (fig. S8A). This in-

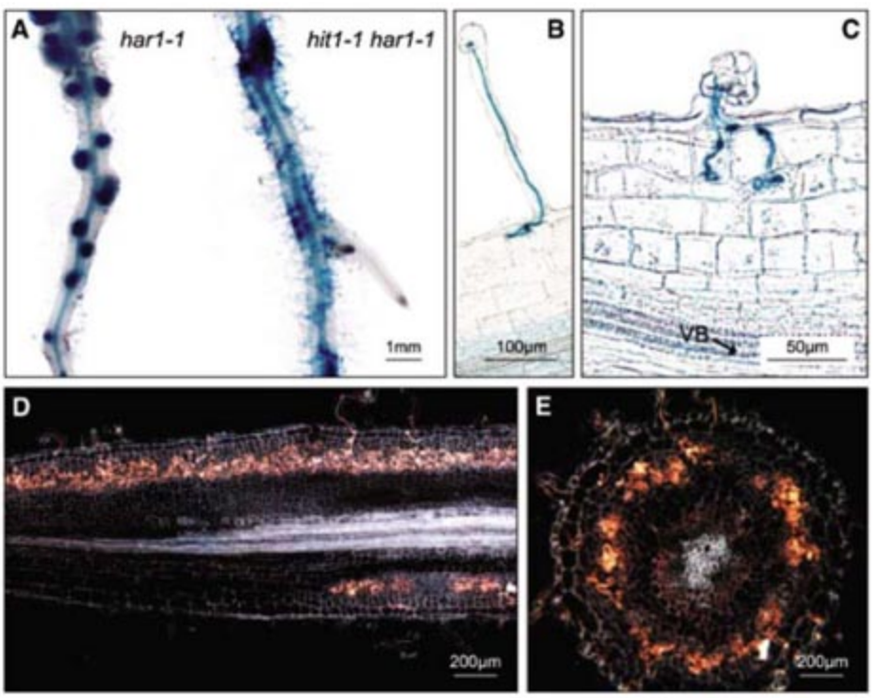


Fig. 1. Root segments stained for β-galactosidase (*LacZ*) activity 10 dai [(A) to (C)] and 14 dai [(D) and (E)] with *M. loti*. (**A**) A large number of ITs gave a blue appearance to the *hit1-1 har1-1* root. (**B**) An IT (blue) traversed a root hair but became blocked at the interface between the epidermis and cortex. (**C**) Misguided ITs looping within the root cortex. The subepidermal infection would normally be associated with subtended cell divisions (*28*). VB, vascular bundle. (**D** and **E**) Negative images of (D) a longitudinal section through *hit1-1 har1-1* root showing a large number of ITs (red) within the root cortex and (E) a cross-section of the *hit1-1 har1-1* root showing accumulation of ITs around the entire midcortex perimeter of the section plane.

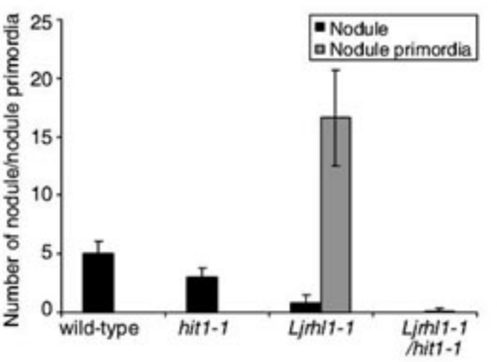


Fig. 2. Numbers of nodules and nodule primordia (10 dai) on wild-type roots and *hit1-1*, *Ljrhl1-1*, and *Ljrhl1-1 hit1-1* mutants. Error bars show means ± 95% CI (n = 20).

duction required de novo protein biosynthesis (fig. S8B). In *hit1-1* roots, BA-stimulated accumulation of *Nin* transcripts was significantly diminished in comparison with wild-type roots, indicating that the high level of *Nin* expression requires the functional *Lhk1* (Fig. 4B).

NF signaling regulates *Nin* expression (*3*, *15*), which is required for the formation of ITs in the root epidermis and initiation of nodule primordia organogenesis in the root cortex (*12*). Our data indicate that although necessary for *Nin* expression and nodule organogenesis, *Lhk1* is not required for IT formation (Fig. 4C). The reported partitioning of *Nin* expression between the root epidermis and cortex could provide a plausible explanation for this apparent conundrum (*3*, *15*, *26*). *Nin* expression, supporting IT formation, may be regulated by an *Lhk1*-independent mechanism in the root epidermis, possibly involving another cytokinin receptor. Diminished nodule organogenesis in *hit1-1 har1-1* and *hit1-1* likely restricts local and/or systemic feedback mechanisms that limit root susceptibility to *Rhizobium* infection, resulting in hyperinfection (Fig. 4C).

The LHK1 homologs, such as LHK2 and LHK3, are likely to function as cytokinin receptors, which may explain a leaky (formation of some nodules) symbiotic phenotype and lack of more general developmental abnormalities in mutants carrying *hit1* alleles. The *snf2* mutant described in the accompanying manuscript (*23*) and strongly reduced nodulation in *Medicago truncatula* plants carrying a *MtCRE1* silencing construct (*27*) further demonstrate that cytokinin sensing is required to stimulate nodule development. Together, these results specify that the regulators of cytokinin biosynthesis and/or action are crucial downstream targets of NF perception (Fig. 4C) and that recruitment of a cytokinin receptor could have been an essential event during the evolution of nitrogen-fixing nodule symbiosis.

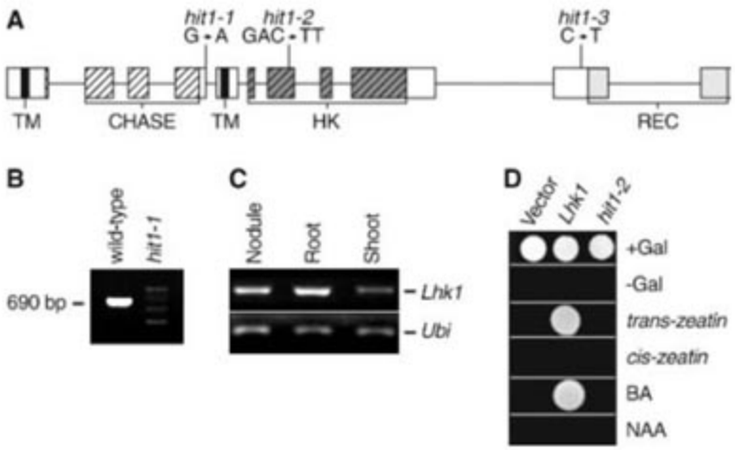


Fig. 3. (**A**) The exon (box)/intron (line) structure of the *Lhk1* gene with positions of the molecular lesions for each of the *hit1* mutant alleles indicated. The corresponding protein domains are indicated: transmembrane domains (TM) (black); CHASE domain (white hatched); histidine kinase (HK) domain (gray hatched); receiver (REC) domain (gray). *hit1-1* carries a nucleotide substitution from G_{1695} to A that leads to alternatively spliced products, as shown by RT-PCR in (**B**). *hit1-2* contained two consecutive transversions (G_{2407} and A_{2408} to TT) followed by a single base (C_{2409}) deletion. This frame shift results in a premature stop codon 45 base pairs (bp) downstream. *hit1-3* has a transition from C_{4922} to T resulting in a premature stop codon before the REC domain. (**C**) Expression of *Lhk1* in various *L. japonicus* tissues as assayed by RT-PCR. Ubiquitin (*Ubi*) was used as the RNA loading control. (**D**) *Lhk1* cDNA confers cytokinin responsiveness to yeast cells. TM182 cells were transformed with either p415CYC carrying the *Lhk1* cDNA, the *hit1-2* cDNA or the vector alone. Transformants were plated on a minimal medium with or without galactose (gal), or on a galactose-free medium supplemented with the indicated hormone. NAA, 1-naphthaleneacetic acid

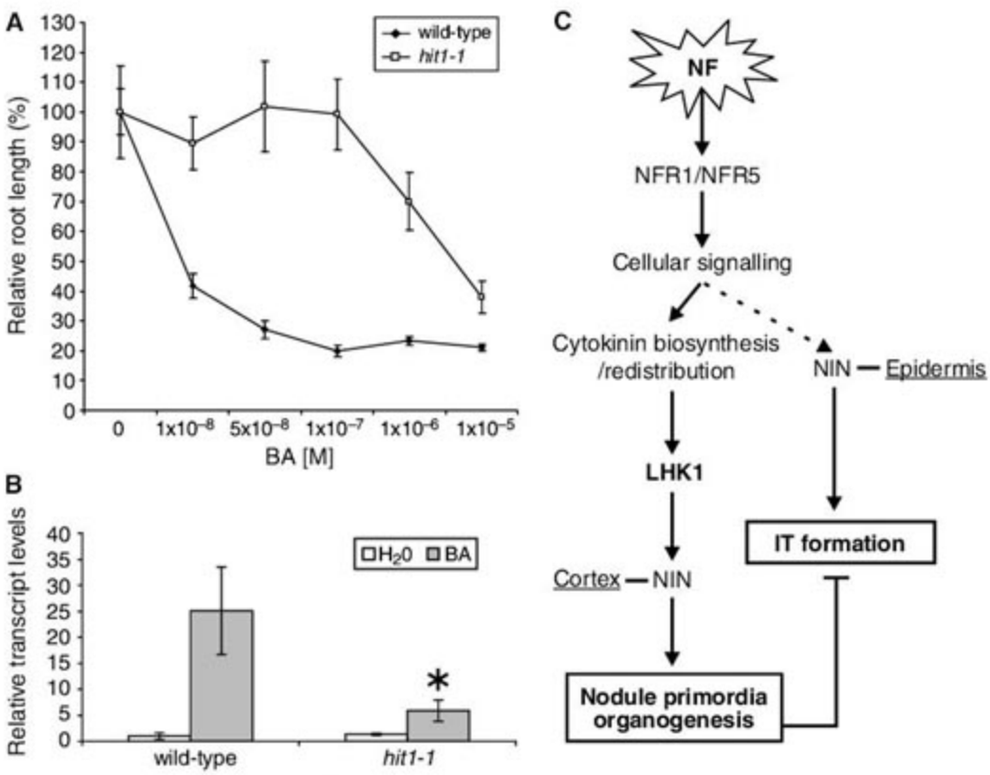


Fig. 4. (**A**) *hit1-1* roots are insensitive to exogenously applied cytokinin. (**B**) qRT-PCR showing significant attenuation (*$P < 0.05$) of *Nin* mRNA in *hit1-1* versus wild-type roots upon exogenous application of 50 nM BA. BA-stimulated accumulation of *Nin* transcripts in *hit1* roots was significantly higher ($P < 0.05$) in comparison to water treatments. Error bars show means ± 95% CI. (**C**) The proposed role of LHK1 in NF-induced nodule organogenesis. Perception of NF by a presumed NFR1 and NFR5 receptor complex stimulates local cytokinin biosynthesis in, or redistribution of cytokinin to, the root epidermis and cortex (*29*). This is perceived by LHK1, which activates *Nin* expression in the root cortex leading to initiation of nodule organogenesis.

References and Notes

1. G. E. D. Oldroyd, M. J. Harrison, M. Udvardi, *Plant Physiol.* **137**, 1205 (2005).
2. E. B. Madsen *et al.*, *Nature* **425**, 637 (2003).
3. S. Radutoiu *et al.*, *Nature* **425**, 585 (2003).
4. E. Limpens *et al.*, *Science* **302**, 630 (2003).
5. S. Stracke *et al.*, *Nature* **417**, 959 (2002).
6. H. Imaizumi-Anraku *et al.*, *Nature* **433**, 527 (2005).
7. J. M. Ané *et al.*, *Science* **303**, 1364 (2004).
8. N. Kanamori *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 359 (2006).
9. J. Lévy *et al.*, *Science* **303**, 1361 (2004).
10. P. Smit *et al.*, *Science* **308**, 1789 (2005).
11. P. Kalo *et al.*, *Science* **308**, 1786 (2005).
12. L. Schauser, A. Roussis, J. Stiller, J. Stougaard, *Nature* **402**, 191 (1999).
13. L. Tirichine, E. K. James, N. Sandal, J. Stougaard, *Mol. Plant Microbe Interact.* **19**, 373 (2006).
14. C. Gleason *et al.*, *Nature* **441**, 1149 (2006).
15. L. Tirichine *et al.*, *Nature* **441**, 1153 (2006).
16. J. B. Cooper, R. L. Long, *Plant Cell* **6**, 215 (1994).
17. A. M. Hirsch, T. V. Bhuvaneswari, J. G. Torrey, T. Bisseling, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 1244 (1989).
18. J. Murray *et al.*, *Mol. Plant Microbe Interact.* **19**, 1082 (2006).
19. L. Krusell *et al.*, *Nature* **420**, 422 (2002).
20. E. Minami *et al.*, *Plant J.* **10**, 23 (1996).
21. B. Karas *et al.*, *Plant Physiol.* **137**, 1331 (2005).
22. C. Kistner, M. Parniske, *Trends Plant Sci.* **7**, 511 (2002).
23. L. Tirichine *et al.*, *Science* **315**, 104 (2007); published online 16 November 2006 (10.1126/science.1132397).
24. T. Maeda, S. M. Wurgler-Murphy, H. Saito, *Nature* **369**, 242 (1994).
25. Y. Fang, A. M. Hirsch, *Plant Physiol.* **116**, 53 (1998).
26. M. Gronlund *et al.*, *Mol. Plant Microbe Interact.* **18**, 414 (2005).

27. S. Gonzalez-Rizzo, M. Crespi, F. Frugier, *Plant Cell* **18**, 2680 (2006).
28. K. Szczyglowski *et al.*, *Mol. Plant Microbe Interact.* **11**, 684 (1998).
29. D. P. Lohar *et al.*, *Plant J.* **38**, 203 (2004).
30. We thank A. Molnar for preparation of figures, L. Ross for technical help, and S. Kosuta and M. Held for helpful comments on the manuscript. This work was funded by Agriculture and Agri-Food Canada Crop Genomics Initiative and National Sciences and Engineering Research Council of Canada grant no. 3277A01 to K.S.

Supporting Online Material
www.sciencemag.org/cgi/content/full/1132514/DC1
Materials and Methods
Figs. S1 to S8
References and Notes

14 July 2006; accepted 27 October 2006
Published online 16 November 2006;
10.1126/science.1132514
Include this information when citing this paper.

A Gain-of-Function Mutation in a Cytokinin Receptor Triggers Spontaneous Root Nodule Organogenesis

Leïla Tirichine,[1] Niels Sandal,[1] Lene H. Madsen,[1] Simona Radutoiu,[1] Anita S. Albrektsen,[1] Shusei Sato,[2] Erika Asamizu,[2] Satoshi Tabata,[2] Jens Stougaard[1]*

Legume root nodules originate from differentiated cortical cells that reenter the cell cycle and form organ primordia. We show that perception of the phytohormone cytokinin is a key element in this switch. Mutation of a *Lotus japonicus* cytokinin receptor gene leads to spontaneous development of root nodules in the absence of rhizobia or rhizobial signal molecules. The mutant histidine kinase receptor has cytokinin-independent activity and activates an *Escherichia coli* two-component phosphorelay system in vivo. Mutant analysis shows that cytokinin signaling is required for cell divisions that initiate nodule development and defines an autoregulated process where cytokinin induction of nodule stem cells is controlled by shoots.

Differentiated plant cells have an unusual capacity for rejuvenating by dedifferentiation and subsequent differentiation to form new organs or complete plants. In the model legume *Lotus japonicus* (lotus), nodule organogenesis is initiated by dedifferentiation of root cortical cells followed by cell proliferation, establishing a cluster of meristematic cells that give rise to the nodule primordium. The developmental process is triggered by compatible *Mesorhizobium loti* bacteria synthesizing lipochitin-oligosaccharide nodulation factor (Nod factor) acting as a mitogen and/or morphogen when recognized by the host plant Nod factor receptors, NFR1 and NFR5 (*1*, *2*). Bacterial invasion of primordia occurs via infection threads progressing through root hairs into the root cortex. Ultimately, rhizobia released from infection threads are endocytosed into cells, which become the infected nitrogen-fixing nodule cells. At the same time, pattern formation and cell differentiation specify tissue and cell types of the new specialized organ, which in turn supplies the plant with nitrogen fixed by endocytosed bacteria.

To dissect the genetic regulation of cellular dedifferentiation and meristem formation, we isolated lotus mutants developing root nodules spontaneously. The *snf2* (spontaneous nodule formation) mutants develop white rhizobia-free nodules in the absence of *M. loti* (Fig. 1, A and B). Detailed histological analysis of nodule sections demonstrates that spontaneous nodules are genuine nodules with an ontogeny and physiology similar to rhizobially induced nodules (*3*). The *snf2* allele is monogenic dominant, and inoculation of *snf2* mutants with *M. loti* results in development of normal nitrogen-fixing root nodules, which strongly suggests the presence of a gain-of-function mutation in this allele. Genetic mapping of *snf2* and sequencing of bacterial artificial chromosome clones identified a homolog of *Arabidopsis* histidine kinase genes (*AHK*) encoding cytokinin receptor proteins (fig. S1A). In light of physiological studies on phytohormones in nodulation (*4*, *5*), this histidine kinase was a likely candidate gene, and the corresponding gene region of *snf2* was sequenced. A single nucleotide transition (C to T), resulting in replacement of a conserved leucine 266 by phenylalanine (L266F), identifies *snf2* as an allele of a lotus histidine kinase (*Lhk1*) gene. Alignment of genomic and cDNA sequences defined a primary structure of *Lhk1* consisting of 11 exons (fig. S1B). Steady-state levels of *Lhk1* transcripts in different plant organs were determined by quantitative reverse transcription polymerase chain reaction. *Lhk1* was expressed at the highest level in roots, nodules, and leaves, but transcripts were present in all organs tested. (Fig. 2A).

Constructs carrying either the *snf2* mutant gene or the wild-type *Lhk1* gene were transformed into wild-type roots using *Agrobacterium rhizogenes*. To assure reliable transfer to transgenic roots of the gene constructs used throughout this study, they were integrated directly into *A. rhizogenes* transferred DNA (T-DNA) by using a recombination approach (*6*, *7*). Thus, they were transformed into plant cells together with the T-DNA, which gave rise to transgenic roots at the hypocotyl wound site. An *Lhk1* gene segment and the corresponding *snf2* gene segment were introduced, and nodulation was scored in the absence of rhizobia. Spontaneous nodulation was observed on transgenic roots transformed with the *snf2* construct, whereas the *Lhk1* wild-type gene was unable to confer spontaneous nodulation (table S1 and fig. S2). This differential response illustrates the effect of the dominant *snf2* mutation and confirms that spontaneous nodulation is caused by a single amino acid substitution in the cytokinin receptor. The absence of nodules on the normal root systems, which served as internal controls for the *A. rhizogenes*–induced *snf2* transgenic roots, and the lack of rhizobia in the nodules that were formed on the *snf2* transgenic roots show that they were indeed spontaneously formed nodules.

An open reading frame of 2979 nucleotides is predicted in the *Lhk1* cDNA clone. The conceptual cytokinin receptor protein (LHK1) consists of 993 amino acids (Fig. 3). At the N terminus, two membrane-spanning segments are located between amino acids 37 and 57 and between amino acids 328 and 357. Located between these segments are motifs characteristic of cytokinin-binding (CHASE) domains. This predicted extracellular domain is followed by a putative intracellular histidine kinase and a receiver domain. These domains are characteristic of two-component regulatory systems operating through phosphorelay.

Comparative analysis defines LHK1 as a member of the cytokinin receptor family (fig. S3). Among the three *Arabidopsis* cytokinin receptors, LHK1 has 68% identity to AHK4/(Cre1), which is important for normal root development and serves a function in perception of externally supplied cytokinin (*8*). The leucine 266 replaced by a phenylalanine in the *snf2* allele is part of a conserved motif shared among the extracellular CHASE domains of histidine kinase receptors (fig. S3). Spontaneous nodulation resulting from an amino acid change located in the CHASE domain suggested a cytokinin-independent function caused by the L266F substitution. To test this hypothesis, we assayed the in vivo activity of lotus wild-type and gain-of-function receptors using the two-component phosphorelay assay developed in *E. coli* (*9*). Functional expression of a cytokinin receptor in an *E. coli* strain lacking the RcsC sensor, which normally regulates extracellular polysaccharide synthesis, allows cytokinin perception to be read out as β-galactosidase activity from a

[1]Laboratory of Gene Expression, Department of Molecular Biology, University of Aarhus, Gustav Wieds Vej 10, DK-8000 Aarhus C, Denmark. [2]Kazusa DNA Research Institute, Kisarazu, Chiba, 292-0818, Japan.

*To whom correspondence should be addressed. E-mail: stougaard@mb.au.dk

cps::lacZ fusion. Expression of the mutant L266F protein does indeed induce β-galactosidase activity in the absence of cytokinin (Fig. 4A). In contrast, wild-type LHK1 induced β-galactosidase activity in a cytokinin-dependent fashion (Fig. 4A). Quantitative determination of β-galactosidase activity in *E. coli* cultures shows that L266F-expressing cells have three times the β-galactosidase activity that control cells and cells expressing wild-type LHK1 have (Fig. 4B). Cytokinin addition results in a twofold induction of β-galactosidase activity in LHK1 cells, whereas L266F cells respond with only a marginal increase in activity. These results demonstrate that LHK1 is a cytokinin receptor and that the L266F receptor exhibits cytokinin-independent activity at a level comparable to the cytokinin-induced activity of the wild-type receptor. We propose that the extracellular CHASE domain, normally binding cytokinin to activate the kinase (*10*–*12*), in the L266F mutant receptor is locked within an active conformation. This hypothesis would explain both the genetic dominant nature of the *snf2* allele and the phosphorelay assay results.

Arabidopsis ahk2 ahk3 ahk4 triple mutants display a reduced number of root vascular cell files, because periclinal procambial cell divisions are impaired (*13*). *snf2* mutant roots have the opposite phenotype with extra layers (Fig. 1, C and D). In explants cultivated without phytohormones (Fig. 1, I and J), cell proliferation was even more pronounced. Additional cell layers originating from periclinal divisions were observed, together with an increase in vascular cell numbers (Fig. 1, E and F). In order to examine possible global effects on cell differentiation, we monitored the in vitro performance of *snf2* and wild-type hypocotyl and root explants (Fig. 1 and fig. S4). The overall hormone dose response is similar. However, in line with the cytokinin-independent response, *snf2* explants survive better at high auxin and develop less callus on cytokinin (Fig. 1, G and H).

Cytokinin-induced changes in cellular processes in plants are accompanied by increased expression of type-A response regulator (*ARR*) genes (*14*). Among the type-A genes, *ARR5* is a rapidly induced response gene, and an *Arabidopsis ARR5* promoter fused with the *Gus* (β-glucuronidase) reporter gene (GUS fusion) was

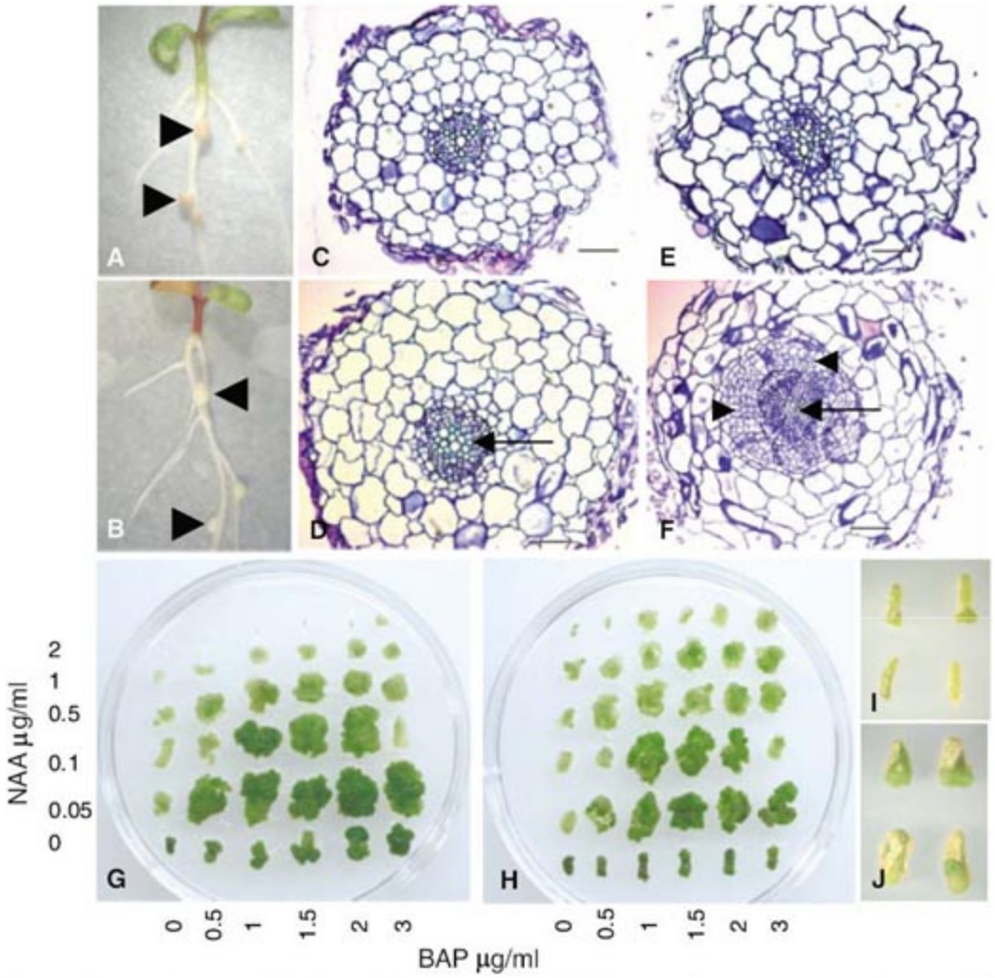
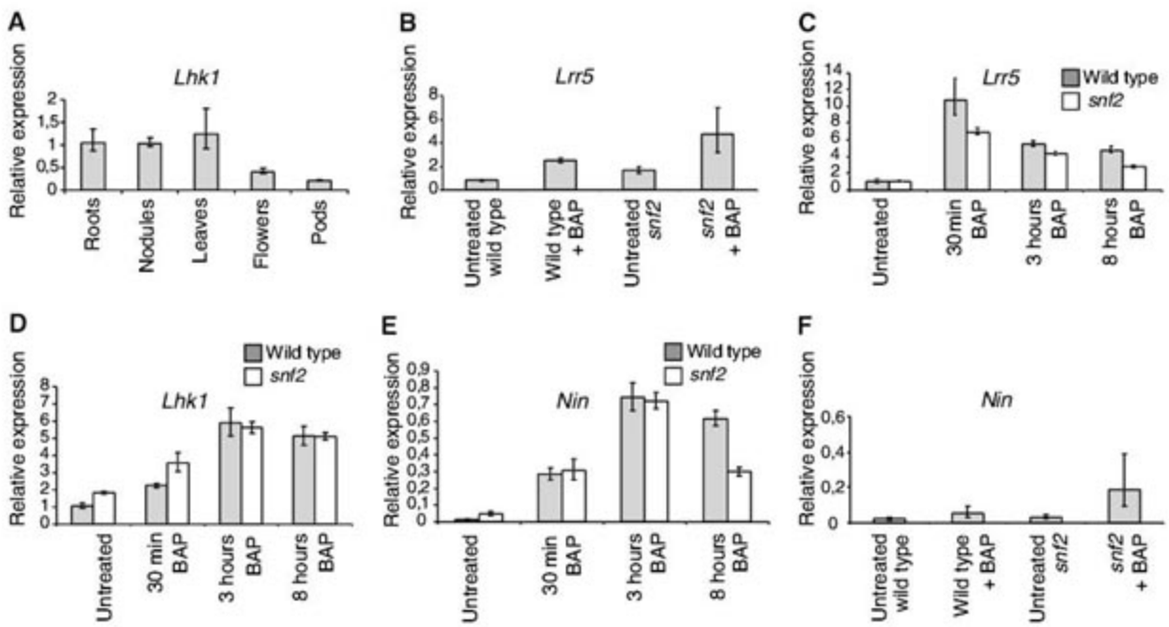


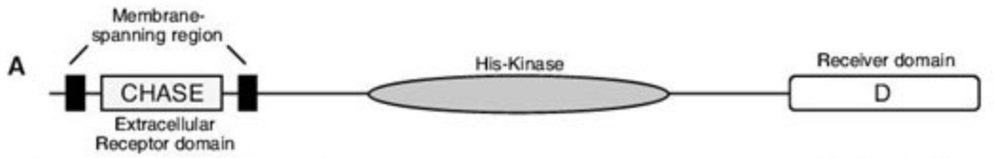
Fig. 1. Phenotypic characterization of the *snf2* mutant. (**A**) Wild-type rhizobia induced root nodule (**B**) spontaneous *snf2* root nodule. Arrowheads, 5-week-old nodules. Transverse section of (**C**) wild-type and (**D**) *snf2* root at time 0 and (**E**) wild-type and (**F**) *snf2* root after 6 days on hormone-free medium. Arrowhead, dividing cells in the pericycle; arrows, xylem cells. (**G**) and (**H**) Callus growth from hypocotyls of wild-type and *snf2* given different concentrations of auxin and cytokinin. Root segments of wild-type (**I**) and *snf2* (**J**) incubated 3 weeks on hormone-free media. Scale bars: (C to F), 50 µm.

Fig. 2. Expression of the *Lhk1* gene in organs and the *Lhk1*, *Lrr5*, and *Nin* in response to cytokinin. (**A**) Expression of *Lhk1* in different organs. (**B**) Expression of *Lrr5* in wild-type and *snf2* root explants incubated on medium with or without 0.5 µg/ml of 6-benzylaminopurine (BAP) for 10 days. (**C** to **E**) Expression of *Lrr5*, *Lhk1*, and *Nin* in intact wild-type and *snf2* roots in response to 10 µM cytokinin. (**F**) Expression of *Nin* in wild-type and *snf2* root explants incubated on medium with or without 0.5 µg/ml of BAP for 10 days.



expressed during nodulation (*15*). Because the L266F receptor protein has cytokinin-independent activity in the *E. coli* assay, we determined transcript levels of a lotus *ARR5* homolog named *Lrr5* (Fig. 2). Following the examples from previous analyses of the complex cytokinin circuitry in *Arabidopsis* (*16*, *17*), we determined transcript levels in both intact plants and in vitro cultivated plant cells in order to capture the dynamics and the range of cytokinin regulation. *Lrr5* transcript in root explants of *snf2* mutants incubated without hormones was found to be two times that of wild-type explants, whereas cytokinin addition increased *Lrr5* transcript level two- to threefold in both (Fig. 2B). Cytokinin treatment of roots increased the *Lrr5* transcript level in *snf2* and wild-type roots, but no difference in expression was detected between untreated *snf2* and wild-type roots (Fig. 2C). Cytokinin also regulates expression of the *Lhk1* gene and induces a rapid increase in *Lhk1* gene transcripts after treatment in both wild-type and *snf2* mutants (Fig. 2D). We also tested whether the lotus *Nin* gene known to be required for initiation of nodule primordia was ectopically expressed in *snf2* roots. As shown in Fig. 2E, the *Nin* gene is up-regulated by cytokinin, and the transcript levels in untreated *snf2* roots were significantly different from those of wild-type roots. No ectopic expression of *Nin* in root explants of *snf2* mutants incubated on hormone-free medium was detected (Fig. 2F). Attenuation of the cytokinin response pathway, as previously described in *Arabidopsis* exposed to cytokinin (*16*, *17*), was also observed in lotus wild-type roots and was even more pronounced in *snf2* (Fig. 2, C and E). Thirty minutes after exposure to exogenous cytokinin, a sevenfold increase in the steady-state level of *Lrr5* transcript was detected in *snf2* roots. In spite of the continuous presence of cytokinin, this initial induction was attenuated at later time points, and only a twofold increase in transcript level was found after 8 hours. In *Arabidopsis*, attenuation of responses to cytokinin is mediated by a complex feedback mechanism. The cytokinin oxidases, which by themselves are cytokinin inducible, and a range of negatively acting response regulators, including ARR5, which also negatively autoregulate their own transcription, were shown to be involved (*16*, *17*). In the gain-of-function *snf2*, the cytokinin hypersensitivity (fig. S5) and the presence of the LHK2 and LHK3 receptors (*18*), which remain cytokinin-dependent, appear to reset the balance point of negative regulation at a level where the transcriptional up-regulation of *Lrr5* in untreated *snf2* roots is relatively small or undetectable. Although transcriptional changes in the *snf2* mutants were limited, plant growth is strongly affected by external cytokinin. In line with the cytokinin-independent activity of the gain-of-function receptor observed in the *E. coli* assay (Fig. 4A) and in the in vitro culture experiments (Fig. 1, H and J; Fig. 2, B, D, and E), *snf2* shoot and root growth was hypersensitive to cytokinin (fig. S5). Prolonged exposure of wild-type plants (8 weeks) to lower cytokinin levels than those used in the experiment shown in fig. S5 did lead to development of small "bumps" that resembled nodule primordia.

The phenotype of *snf2* mutants suggests that cytokinin signaling acts downstream of Nod-factor signal transduction. To test this hypothesis, the *snf2* gene construct was transformed into mutants of the Nod-factor signal transduction pathway and in mutants impaired in downstream genes. *snf2*-mediated spontaneous nodulation in *nfr1-1*, *nfr5-2* Nod-factor receptor single and double mutants lacking the earliest electrophysiological responses (*1*, *2*) demonstrates a function for *Lhk1* downstream of Nod-factor signal perception. The signal transduction *symRK* mutants lacking Ca^{2+} spiking (*19*) and *ccamk* (Ca^{2+}- and calmodulin-dependent protein kinase) mutants, which have been suggested to be unable to interpret Ca^{2+} spiking (*20*, *21*), also develop spontaneous nodules in *snf2* transgenic roots. Incidentally, these results provide independent evidence for *snf2*-mediated spontaneous nodula-



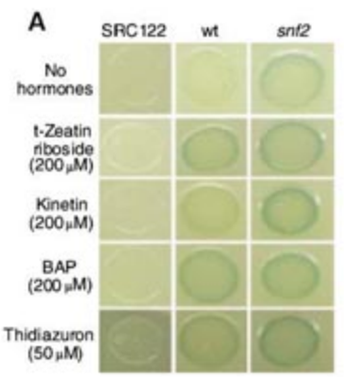
B MGLGFKMQQSHHPVALKLHEQAGSQRKFTFIQNFRN
WFLPLLFVWFIVMAAFGACIY
HKMDAETKVRRKEVLGSLCDQRARMLQDQFSVSVNHVHALAILVSTFHYYRNTSAIDQETFAEYTARTAFERPLMSGVAY AQRVVHSERERFEEKQHGWVIKTMERVPSGVRDEYAAVIFAQETVSYLESIDMMSGEEDRENILRARATGKAVLTSPFRLL DSHHLGVVLTFPVYKSKLPPEPTTEEVIKAIAGYIGGSFDVESLVENLLGQLAGNQAILVKVYDITNSSDPLI

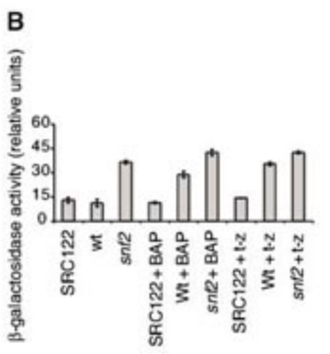
MYGSQYEEGDMSLVHESKLDFGDPYRKHHMICRYHQQ
APTNWIAYTTAFLFFVILCLVGYILYAAGT
HIVKVEDDYNAMQDLKVKAEAA**DIAKSQFLATVSHEIRTPMNGILGMLGLLLRTELSSTQRDYAQTAQACGKA LIALINEV**LDRAKIEAGKLELEAVPFDLRSILDDVLSLFSEKSRHKGLELAVFVSDKVPDIVMGDPGRFRQIVTNLVGNS VKFTERGHIFVKVHLAEKRQCTMNGKCE**TFLNGGCDDVLHVSGSYNLKTLSGYEAADERNSWDNFKHHIADEEF FFDASVKKLASSESYEQVTLMVSVEDTGIGISFSAQDSIFMPFVQADSSTSRNYGGTGIGLSISKCLVELMGGQ INFISRPQVGSTFSFTA**DFGTFKKNSTTDMKKLNFEDLPSSFRGLKAIVVDGKPVRAAVTRYHLKRLGIQAKVAISINK AVSLCGKNGSLTSALFQPDIIFVEKDSWVSGEDGGIFNAFKMPQMILLATNICNAEFDKAKAAGFSDTVIMKPLRASMLA ACLQQVFGTGKTRQFGKDMSNGSSVRSLLCGKK*ILVVDDNLVNRRVAAGALKNFGADVKCAASGKAALEMLQYPHD FDACFMDIQMPEMDGFEATRRIRMMEREASEQLKSESGEENGKKSEFHMPIIAMTADVIHATYDKCLNCGMDGYV SKPFEEENLYQAVAKFFK**SKPASDS

Membrane-spanning domain

Cytoplasmic domain
Membrane-spanning domain

Fig. 3. Structure of the *Lotus* LHK1 protein. (**A**) Schematic representation of the LHK1 protein domains. (**B**) The amino acid sequence of LHK1 arranged in protein domains. The extracellular receptor domain is in italics. The predicted CHASE domain within the extracellular receptor domain is underlined. The asterisk marks the amino acid substitution in the mutant. The histidine kinase domain is bold and underlined. The histidine kinase adenosine triphosphatase (ATPase) domain is bold. The receiver domain is bold and italics.





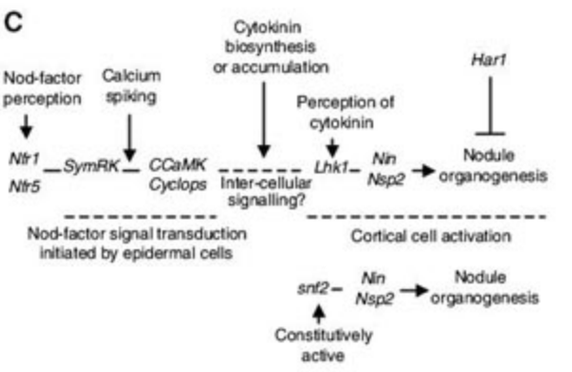


Fig. 4. In vivo assays of receptor-mediated cytokinin signaling. (**A**) Plate assay of β-galactosidase activity expressed from a *cps::lacZ* reporter gene in *E. coli*. The SRC122 strain carrying the *cps::lacZ* reporter transformed with either the gain-of-function *snf2* or *Lhk1* expression construct was grown on plates in the absence or presence of cytokinins. The blue color shows β-galactosidase activity (**B**) Cytokinin induced β-galactosidase activity in liquid cultures of SRC122 *cps::lacZ* transformed with either *snf2* or wild-type constructs. T-z, *trans*-zeatin. (**C**) Model for functional role of *Lhk1* in nodulation. Recognition of Nod-factor by NFR1 and NFR5 induces Nod-factor signal transduction, including calcium spiking and CCaMK kinase activity. A localized increase in cytokinin levels perceived by the LHK1 receptor then leads to cortical cell dedifferentiation and cell cycle activation. *snf2* acts independently of cytokinin but still requires *Nin*, *Nsp2* genes for nodule organogenesis.

tion. The *nfr1*, *nfr5*, *symrk*, and *ccamk* mutants are unable to form nodules in response to rhizobia inoculation. Thus, nodule formation on the *snf2* transgenic roots could not have resulted from contaminating rhizobia. In *nin* and *nsp2* mutants arrested before initiation of cell division induced by Nod-factor signaling, no spontaneous nodules were observed in *snf2* transgenic roots. Because *A. rhizogenes*–induced roots only develop when the hypocotyl wound site infection is used in lotus and because the *snf2* gene construct was integrated into the T-DNA, these results show that cytokinin signal perception acts upstream of cell division initiation (table S2). Furthermore, evidence for a central role of cytokinin and cytokinin perception downstream of Nod-factor signal transduction comes from the additive effect of *snf1-1* and *snf2* mutations. The *snf1-1* mutants synthesize a CCaMK protein impaired in autophosphorylation (*20*, *21*) and develop an average of 7 ± 0.9 (95% confidence interval) spontaneous nodules, whereas *snf2* mutants develop 3 ± 0.5. The *snf1-1 snf2* double mutants exceed both with 17 ± 0.9 spontaneous nodules. Parallel signaling cannot be excluded, but more likely, the deregulated signaling in *snf1* results in a local increase in cytokinin levels transcriptionally up-regulating *snf2* (Fig. 2D) and amplifying spontaneous nodulation. The previously reported expression of a *Nin*-GUS promoter fusion in *snf1* nodule primordia and the absence of epidermal expression in *snf1* roots (*20*) further suggest cytokinin signaling is a cortical response. Conversion of cortical cells into nodule stem cells or subsequent organ development seem therefore tightly controlled. We tested this in a hypernodulating *har1-1* mutant (*22*). Homozygous *snf2 har1-1* double mutants developed an average of 14 ± 1.4 spontaneous nodules, whereas *snf2* mutants developed an average of 3 ± 0.5, and *har1-1*, none (fig. S6). This indicates that only a few cells dedifferentiate or that only a few dedifferentiated cells sustain cell divisions during the *snf2* nodule-initiation process. The shoot controlled autoregulation of the root nodule number (*22*) is thus acting downstream of cytokinin signaling–induced activation of root nodule founder cells (Fig. 4C).

From *Arabidopsis* and tobacco, there is evidence for cytokinin regulation of cell cycle phase transitions (*23*) and for overlapping roles for three AHK receptors in maintaining stem cells and cell divisions during organ formation (*13*). Phytohormones have also been implicated in nodule organogenesis. Applications of auxin transport inhibitors resulted in empty nodule-like structures, which suggested that local inhibition of auxin transport (*24*) sensitizes cells for division. Other experiments showed that externally supplied cytokinin induces cortical cell division and activation of *Enod12*, *Enod40*, and *Enod2* genes (*4*, *25*), and expression of a cytokinin biosynthesis *tzs* gene in a nodulation-deficient *Sinorhizobium meliloti* resulted in nodule-like structures (*5*).

Here we show conclusively that cytokinin signaling plays an important role in plant meristem formation and is directly involved in initiating root nodule organogenesis. The opposite phenotype effects of the *snf2* gain-of-function and *hit1* loss-of-function mutations reported in the accompanying paper (*18*), together with the reduced nodulation observed after down-regulation of the corresponding gene in *Medicago* (*26*), clearly demonstrate that cytokinin signaling is necessary and sufficient for the dedifferentiation and cell proliferation leading to root nodule formation.

References and Notes

1. S. Radutoiu *et al.*, *Nature* **425**, 585 (2003).
2. E. B. Madsen *et al.*, *Nature* **425**, 637 (2003).
3. L. Tirichine, E. K. James, N. Sandal, J. Stougaard, *Mol. Plant Microbe Interact.* **19**, 373 (2006).
4. Y. Fang, A. M. Hirsch, *Plant Physiol.* **116**, 53 (1998).
5. J. B. Cooper, S. R. Long, *Plant Cell* **6**, 215 (1994).
6. J. Stougaard, D. Abildsten, K. A. Marcker, *Mol. Gen. Genet.* **207**, 251 (1987).
7. J. Hansen *et al.*, *Plant Cell Rep.* **8**, 12 (1989).
8. A. P. Mahonen *et al.*, *Genes Dev.* **14**, 2938 (2000).
9. T. Suzuki *et al.*, *Plant Cell Physiol.* **42**, 107 (2001).
10. T. Kakimoto, *Plant Cell Physiol.* **42**, 677 (2001).
11. V. Anantharaman, L. Aravind, *Trends Biochem. Sci.* **26**, 579 (2001).
12. J. Pas, M. von Grotthuss, L. S. Wyrwics, L. Rychlewski, J. Barciszewski, *FEBS Lett.* **576**, 287 (2004).
13. N. Nishimura *et al.*, *Plant Cell Physiol.* **45**, 1485 (2004).
14. C. E. Hutchison, J. J. Kieber, *Plant Cell* **14** (suppl.), S47 (2002).
15. D. P. Lohar *et al.*, *Plant J.* **38**, 203 (2004).
16. A. M. Rashotte, S. D. Carson, J. P. To, J. J. Kieber, *Plant Physiol.* **132**, 1998 (2003).
17. I. Hwang, J. Shen, *Nature* **413**, 383 (2001).
18. J. D. Murray *et al.*, *Science* **315**, 101 (2007); published online 16 November 2006 (10.1126/science.1132514).
19. S. Niwa *et al.*, *Mol. Plant Microbe Interact.* **14**, 848 (2001).
20. L. Tirichine *et al.*, *Nature* **441**, 1153 (2006).
21. C. Gleason *et al.*, *Nature* **441**, 1149 (2006).
22. L. Krusell *et al.*, *Nature* **420**, 422 (2002).
23. T. Kakimoto, *Annu. Rev. Plant Biol.* **54**, 605 (2003).
24. U. Mathesius *et al.*, *Plant J.* **14**, 23 (1998).
25. C. Dehio, F. J. de Bruijn, *Plant J.* **2**, 117 (1992).
26. S. Gonzalez-Rizzo, M. Crespi, F. Frugier, *Plant Cell* **18**, 2680 (2006).

27. L.T. was supported by the Lotus Training Network grant HPRN-CJ-2000-00086. *Lhk1* gene accession number: AM287032; *LHK1* mRNA accession number: AM287033.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1132397/DC1
Materials and Methods
Figs. S1 to S7
Tables S1 and S2
References

12 July 2006; accepted 7 November 2006
Published online 16 November 2006;
10.1126/science.1132397
Include this information when citing this paper.

Differential Antigen Processing by Dendritic Cell Subsets in Vivo

Diana Dudziak,[1] Alice O. Kamphorst,[1] Gordon F. Heidkamp,[1] Veit R. Buchholz,[1] Christine Trumpfheller,[2] Sayuri Yamazaki,[2] Cheolho Cheong,[2] Kang Liu,[1] Han-Woong Lee,[3] Chae Gyu Park,[2] Ralph M. Steinman,[2] Michel C. Nussenzweig[1,4]*

Dendritic cells (DCs) process and present self and foreign antigens to induce tolerance or immunity. In vitro models suggest that induction of immunity is controlled by regulating the presentation of antigen, but little is known about how DCs control antigen presentation in vivo. To examine antigen processing and presentation in vivo, we specifically targeted antigens to two major subsets of DCs by using chimeric monoclonal antibodies. Unlike $CD8^+$ DCs that express the cell surface protein CD205, $CD8^-$ DCs, which are positive for the 33D1 antigen, are specialized for presentation on major histocompatibility complex (MHC) class II. This difference in antigen processing is intrinsic to the DC subsets and is associated with increased expression of proteins involved in MHC processing.

Lymphoid organ DCs are composed of distinct subsets (*1*–*5*). In the spleen, two major types of DCs are found: The first is positive for the CD8 marker and the C-type lectin CD205 ($CD8^+DEC205^+$), and the second lacks CD8 but expresses the antigen recognized by the 33D1 monoclonal antibody (mAb) $CD8^-33D1^+$. These subsets reside in different anatomic locations—$CD8^+DEC205^+$ DCs are in the T cell zone, whereas $CD8^-33D1^+$ DCs are in the red pulp and marginal zone—and the two can be further distinguished by a number of surface markers (*4*, *5*) (Fig. 1, A to C, and fig. S1).

$CD8^+DEC205^+$ DCs appear to be specialized for uptake of dying cells and play a unique role in resistance to certain viral infections (*6*–*8*). Notable among the other distinctions between the two cell types is the suggestion that $CD8^+DEC205^+$ DCs are specialized for cross-presentation, which is the ability to process nonreplicating antigens for presentation to T cells by class I molecules of the major histocompatibility complex (MHCI)

[1]Laboratory of Molecular Immunology, The Rockefeller University, New York, NY 10021, USA. [2]Laboratory of Cellular Immunology and Physiology, The Rockefeller University, New York, NY 10021, USA. [3]Department of Biochemistry, College of Sciences, Yonsei University, Seoul 120-749, Korea. [4]Howard Hughes Medical Institute, The Rockefeller University, New York, NY 10021 USA.

*To whom correspondence should be addressed. E-mail: nussen@rockefeller.edu

(*7*–*12*). However, a direct comparison of the capacity of the two subsets to process antigen in vivo has been lacking.

We identified the antigen recognized by the 33D1 mAb through a combination of gene array and candidate gene approaches and found it to be dendritic cell inhibitory receptor–2 (DCIR2) [Fig. 1, D and E; fig. S2 and (*13*)]. DCs are highly enriched in lectins, and in addition to DCIR2 and DEC205, the $CD8^+DEC205^+$ and $CD8^-33D1^+$ DCs differed in expression of a number of other lectins (Fig. 1D and fig. S2B). To evaluate regulation of antigen processing and T cell activation by the two DC subsets in vivo, we delivered antigens to each cell type in situ using chimeric αDEC205 (*14*) and 33D1 antibodies (*15*) [fig. S3 and (*13*)]. Delivery of ovalbumin (OVA) antigen by the injected antibodies was monitored after DC purification by antigen presentation in vitro to transgenic OT-I or OT-II T cells specific for OVA peptides presented on MHC class I or MHC class II, respectively (*14*, *16*). As previously reported, purified DCs targeted with αDEC205-OVA in vivo induced both OT-I and OT-II T cell proliferation in vitro, although the extent of OT-II proliferation was relatively modest (Fig. 1F) (*14*, *16*). In contrast, antigen delivered with the 33D1 antibody elicited high levels of OT-II but no detectable OT-I T cell responses (Fig. 1F). B cells and other non-DCs purified from mice injected with αDEC205-OVA or 33D1-OVA failed to present OVA either to OT-I or OT-II T cells (Fig. 1F). The specificity of targeting was made apparent by the specific localization of αDEC205 on $CD8^+DEC205^+$ DCs and 33D1 on $CD8^-33D1^+$ DCs after chimeric antibody injection in vivo (Fig. 1G). In addition, both antibodies were internalized by the cells, although the kinetics of 33D1 internalization was slower than for DEC205, and the amount of internalized αDEC205-OVA was greater than the amount of internalized 33D1-OVA (Fig. 1H and fig. S4). Neither antibody altered DC maturation status, as determined by surface expression of MHCII, CD40, CD69, CD80, and CD86 (fig. S5). Finally, the difference in presentation between the two DC subsets in vivo was not due to a difference in their ability to present peptides once processed, because the two were equivalent in presentation of antigen to the same transgenic T cells when processed peptides were added to in vitro cultures (fig. S6). We conclude that DCs targeted by αDEC205-OVA or 33D1-OVA in vivo are distinct in their ability to present antigen on MHCI and MHCII in vitro.

To examine T cell activation in response to antigen presentation in vivo, we labeled OT-I and OT-II T cells with 5-(6)-carboxyfluorescein diacetate succinimidyl diester (CFSE), a reporter dye for cell division, and monitored the cells after adoptive transfer to a new host (*13*). As observed in vitro, αDEC205-OVA induced MHC class I–restricted OT-I responses with greatest efficiency, whereas 33D1-OVA primarily elicited MHC class II–restricted OT-II responses

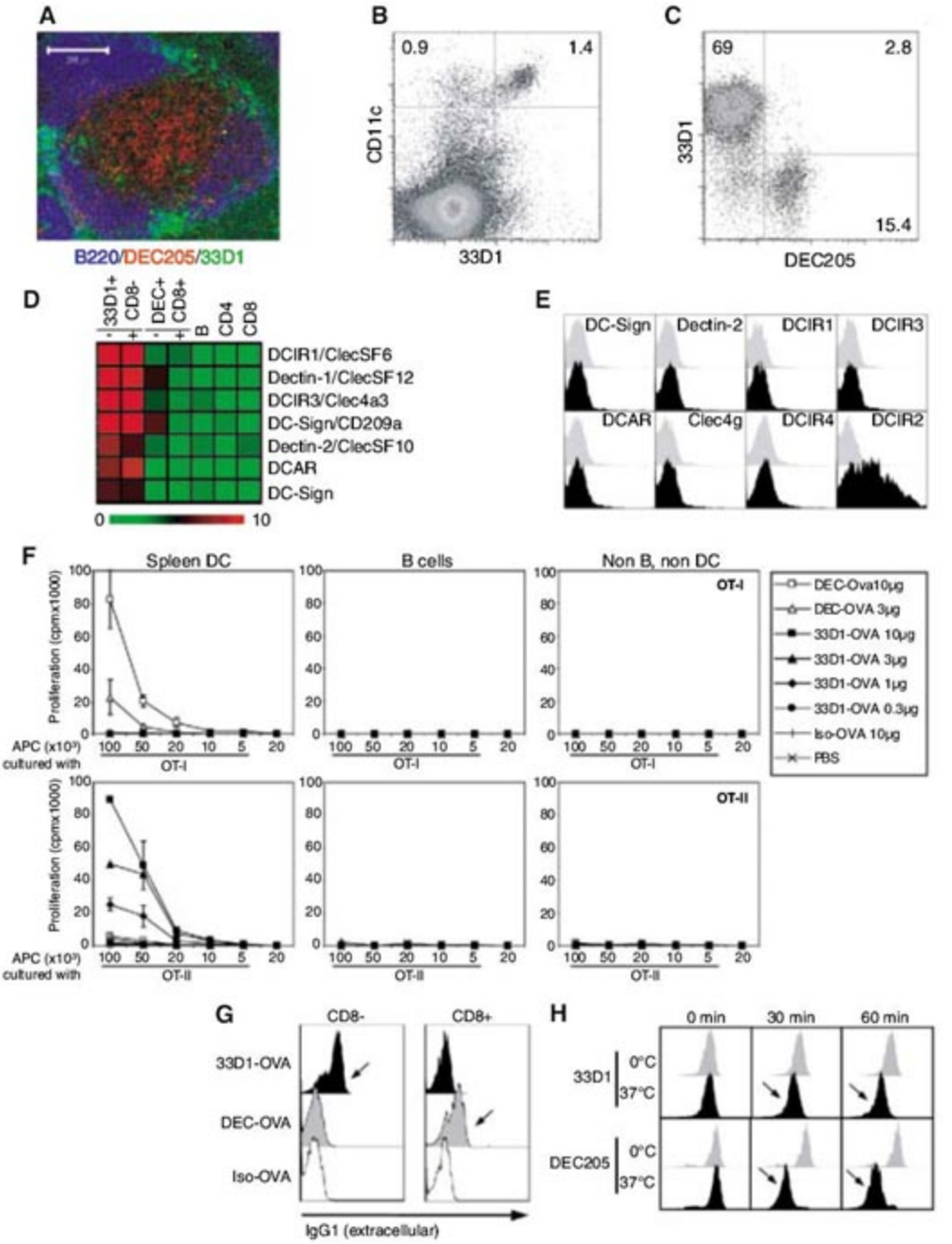


Fig. 1. CD4 and CD8 T cell responses to targeted antigen in vitro. (**A**) Micrograph shows immunohistochemistry of 33D1 (green), αDEC205 (red), and αB220 (blue). (**B** and **C**) Dot plots show splenocytes analyzed by flow cytometry for expression of CD11c and 33D1 (B) and 33D1 and DEC205 gated on $CD11c^{high}$ cells (C). Numbers indicate percentages of total splenocytes and $CD11c^+$ splenocytes, respectively. (**D**) Affymetrix gene array analysis of candidate C-type lectins. RNA was prepared from FACS-sorted B cells, $CD4^+$ T cells, $CD8^+$ T cells, $CD8^+DEC205^+$, and $CD8^-33D1^+$ WT (–) and Flt3L-melanoma (+) injected mice. Each bar represents a mean of three individual gene arrays. List of candidate genes showing difference between $CD8^+DEC205^+$ and $CD8^-33D1^+$ subsets in WT- and Flt3L-injected mice. (**E**) Histograms show 33D1-A647 (black) and isotype IgG2b-A647 antibody (gray) staining of 293T cells transiently transfected with the indicated cDNAs. (**F**) Graphs show [^{3}H]thymidine incorporation by 10^5 OT-I (upper panels) or OT-II T cells (lower panels) cultured with the indicated numbers of DCs, B cells, or non-B non-DCs purified from C57BL/6 mice injected with 10 μg of 33D1-OVA, αDEC-OVA, or Iso-OVA 12 hours earlier. (**G**) Histogram shows extracellular αDEC205 or 33D1 antibodies on $CD8^+DEC205^+$ and $CD8^-33D1^+$ DCs 30 min after intravenous injection of 10 μg of αDEC205-OVA, 33D1-OVA, or Iso-OVA control, visualized with anti-mouse IgG-FITC (fluorescein isothiocyanate). (**H**) Histograms show internalization of purified rat 33D1 (upper panels) or rat αDEC205 antibodies (lower panels) after incubation at 37°C for 0, 30, or 60 min (black), or cell surface expression after incubation on ice for a further 0, 30, or 60 min (gray). Cells were gated on $CD11c^+CD8^-$ and $CD11c^+CD8^+$ DCs.

(Fig. 2A). By in vivo dose-response experiments, comparing cell division and T cell expansion after transfer, 33D1-OVA was a factor of 10 less effective in presentation to OT-I T cells and 10 times as effective in presentation to OT-II T cells as αDEC205-OVA (Fig. 2, A and B). Antigen presentation after a single administration of either αDEC205-OVA (*16*) or 33D1-OVA was long-lasting (Fig. 2C). Thus, OT-I T cells proliferated in response to antigen even when they were transferred 10 days after αDEC205-OVA injection, and OT-II T cells proliferated when cells were transferred up to 5 days after 33D1-OVA injection (Fig. 2C). However, T cells proliferating in response to antigen delivered by targeting antibodies in the steady state were rapidly deleted, and the remaining cells were unresponsive to further stimulation in vitro (Fig. 2, D and E). In contrast, when antigen delivery by 33D1 was combined with DC activation by αCD40, the expanded T cell population persisted and demonstrated strong recall responses to antigen challenge in vitro (Fig. 2, F and G). Thus, antigen delivery to $CD8^-33D1^+$ DCs in vivo results in preferential MHCII-restricted antigen presentation. Nevertheless, delivery of antigens to both DC subsets in the steady state leads to T cell tolerance, whereas targeting in combination with DC maturation by CD40 ligation leads to expansion of T cell clones that remain responsive to antigen.

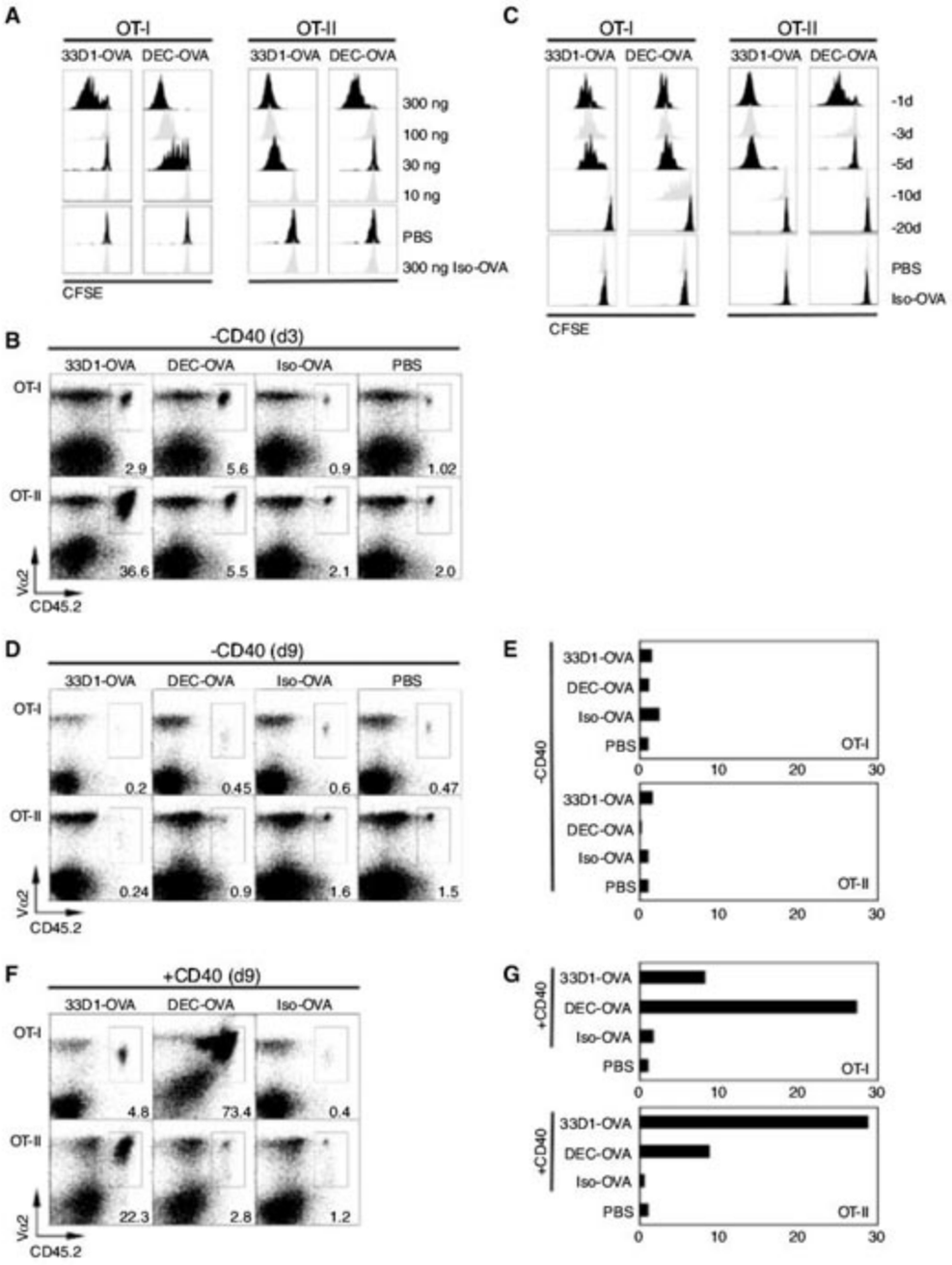


Fig. 2. CD4 and CD8 T cell responses to targeted antigen in vivo. (**A**) Histograms show proliferation as measured by CFSE dye dilution by OT-I (left) or OT-II (right) T cells 3 days after injection of varying amounts of 33D1-OVA, αDEC-OVA, or control antibodies. (**B**) Dot plots show the relative numbers of OT-I (upper panels) or OT-II (lower panels) T cells remaining in spleen 3 days after injection of 300 ng of 33D1-OVA, αDEC-OVA, or Iso-OVA. Numbers indicate percentages of gated $CD4^+$ or $CD8^+$ that were $Va2^+CD45.2^+$. (**C**) As in (A), but T cell transfer was performed 1, 3, 5, 10, or 20 days after injection of 3 μg of 33D1-OVA, αDEC-OVA, or Iso-OVA. (**D**) As in (B), but 9 days after injection of chimeric antibodies. (**E**) Bar graphs show [^{3}H]thymidine incorporation by CD4 or CD8 T cells purified on day 9 after injection of 3 μg of 33D1-OVA, αDEC-OVA, or Iso-OVA and challenged with antigen in vitro. (**F**) Same as (D), except that 50 μg of αCD40 antibody was injected with the targeting antibodies to induce DC maturation. (**G**) Same as (E), except that 50 μg of αCD40 antibody was injected with the targeting antibodies to induce DC maturation. Panels are representative of two to four independent experiments. PBS, phosphate-buffered saline.

To examine the mechanism responsible for differential antigen presentation by the two DC subsets, we assayed formation of MHCII hen egg lysozyme (HEL) peptide complexes (MHCII-p) by using a mAb specific for this complex (*17*). $CD8^+DEC205^+$ DCs showed small amounts of surface MHCII-p 3 hours after injection of αDEC205-HEL, but this was no longer visible after 1 day (Fig. 3A). In contrast, $CD8^-33D1^+$ DCs targeted with 33D1-HEL antibodies displayed much higher levels of MHCII-p after 3 hours and continued to display MHCII-p 2 days after targeting (Fig. 3A, arrows). MHCII-p formation was specific for the targeting antibody and independent of DC activation because mice deficient in the lipopolysaccharide (LPS) Toll-like receptor 4 (TLR-4) were indistinguishable from controls in this assay (Fig. 3, A and B) (*13*). We conclude that antigens delivered by the 33D1 antibody to $CD8^-33D1^+$ DCs in the steady state are processed and transferred to the cell surface as MHCII-p more efficiently than are antigens delivered by αDEC205 to $CD8^+DEC205^+$ DCs.

To determine whether the observed differences in antigen processing were due to cell-intrinsic differences between the two DC subsets, we produced transgenic mice that express human DEC205 (CD11c-hDEC205 B10.BR transgenic mice) on both DC subsets and performed targeting experiments with an αhDEC205 antibody that does not cross-react with mouse DEC205 (*18*) [fig. S7; note that CD11c-hDEC205 mice show position effect variegation and hDEC205 is equally expressed and variegated on both DC subsets (*13*)]. Both subsets were specifically targeted by αhDEC205-HEL in transgenic mice but not wild-type (WT) controls in vivo as measured by surface staining with antibodies to immunoglobulin G (anti-IgG) (fig. S8). However, only $CD8^-33D1^+$ CD11c-hDEC205 transgenic DCs showed high levels of MHCII-p after αhDEC205-HEL or MHCII-p presentation after αhDEC205-OVA injection (Fig. 3C and fig. S9). To further compare presentation by αDEC205-OVA and 33D1-OVA targeting in the same cell, we infected bone marrow–derived DCIR2-negative DCs with a retrovirus encoding DCIR2

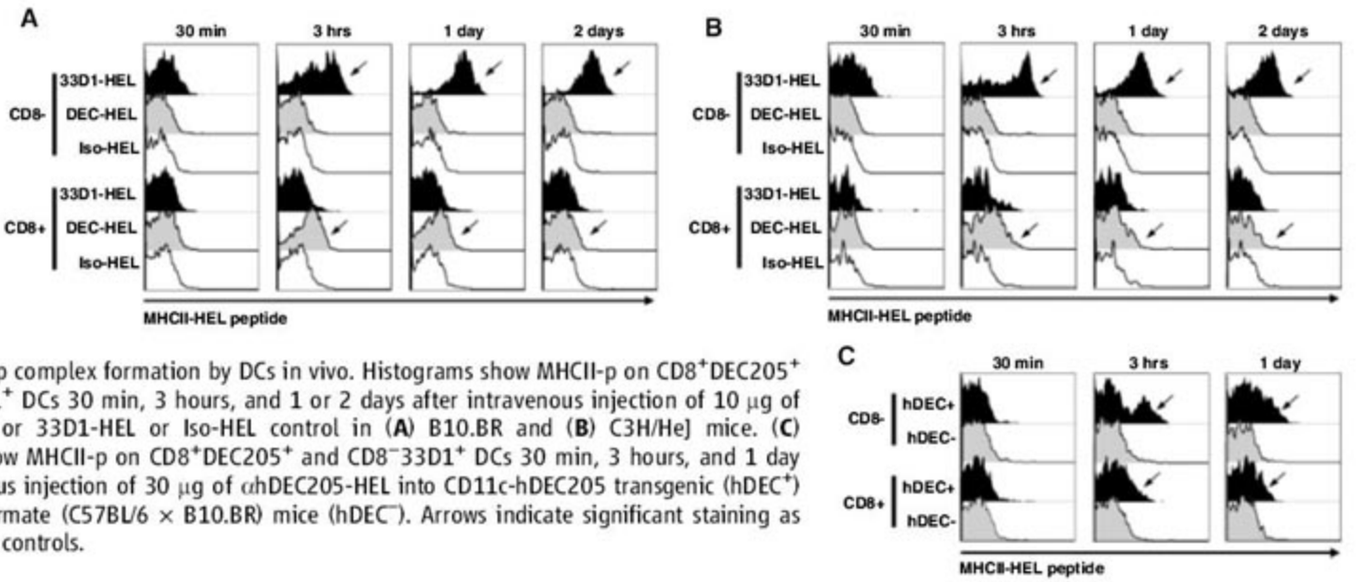


Fig. 3. MHCII-p complex formation by DCs in vivo. Histograms show MHCII-p on $CD8^+DEC205^+$ and $CD8^-33D1^+$ DCs 30 min, 3 hours, and 1 or 2 days after intravenous injection of 10 μg of αDEC205-HEL or 33D1-HEL or Iso-HEL control in (**A**) B10.BR and (**B**) C3H/He] mice. (**C**) Histograms show MHCII-p on $CD8^+DEC205^+$ and $CD8^-33D1^+$ DCs 30 min, 3 hours, and 1 day after intravenous injection of 30 μg of αhDEC205-HEL into CD11c-hDEC205 transgenic ($hDEC^+$) or control littermate (C57BL/6 × B10.BR) mice ($hDEC^-$). Arrows indicate significant staining as compared with controls.

and green fluorescent protein (GFP) (fig. S10A). Infected cells expressing the retrovirally encoded DCIR2 were then sorted on the basis of GFP expression and targeted with αDEC205-OVA or 33D1-OVA in vitro. Following DC maturation with LPS, presentation to OT-II T cells was equivalent for the two targeting antibodies [fig. S10B and (*13*)]. This shows that class II presentation by mature bone marrow–derived DCs (BMDCs) was independent of whether they were targeted by αDEC205-OVA or 33D1-OVA. We conclude that the difference in MHCII processing by DC subsets is an intrinsic property of the cell and not due to differences between the receptors targeted by αDEC205 or 33D1 mAbs.

Many of the proteins that regulate MHCI and MHCII processing pathways have been described and their expression documented in DCs (*19*, *20*). To determine whether the two DC subsets show systematic intrinsic differences in expression of components of the MHCI and MHCII processing machinery, we performed microarray experiments on mRNA isolated from the two DC subsets (Fig. 4, A and B; fig. S11A) (*13*). We found that the two DC subsets differentially express components of the MHCI and MHCII processing pathways in a manner consistent with their ability to produce MHCII-p and induce CD4 and CD8 T cell responses. $CD8^+DEC205^+$ DCs, which are biased for MHCI cross-presentation, were enriched in Tap1, Tap2, calreticulin, calnexin, Sec61, ERp57, ERAAP, as well as cystatin B and C, all of which are involved in MHCI presentation or inhibition of enzymes that process peptides for MHCII presentation (*19*, *20*) (Fig. 4B). In contrast, $CD8^-33D1^+$ DCs, which are biased for MHCII presentation, were enriched in cathepsins C, H, and Z, asparagine endopeptidase (AEP), GILT, and H2-Mbeta 1, all of which are implicated in the MHCII antigen processing pathway (*19*, *20*) (Fig. 4A). To confirm that these proteins were differentially expressed in DC subsets, we performed fluorescence-activated cell sorting (FACS) analysis for intracellular H2DM (Fig. 4C) and Western blotting experiments. MHCI processing–associated proteins were expressed at higher levels in $CD8^+DEC205^+$ DCs (Fig. 4E and fig. S11B), whereas MHCII processing proteins were expressed at higher levels in $CD8^-33D1^+$ DCs (Fig. 4, C to F, and fig. S11B). We conclude that the differences in expression of proteins involved in antigen processing in $CD8^+DEC205^+$ and $CD8^-33D1^+$ DCs are consistent with preferential processing of antigens for presentation by the two cell types.

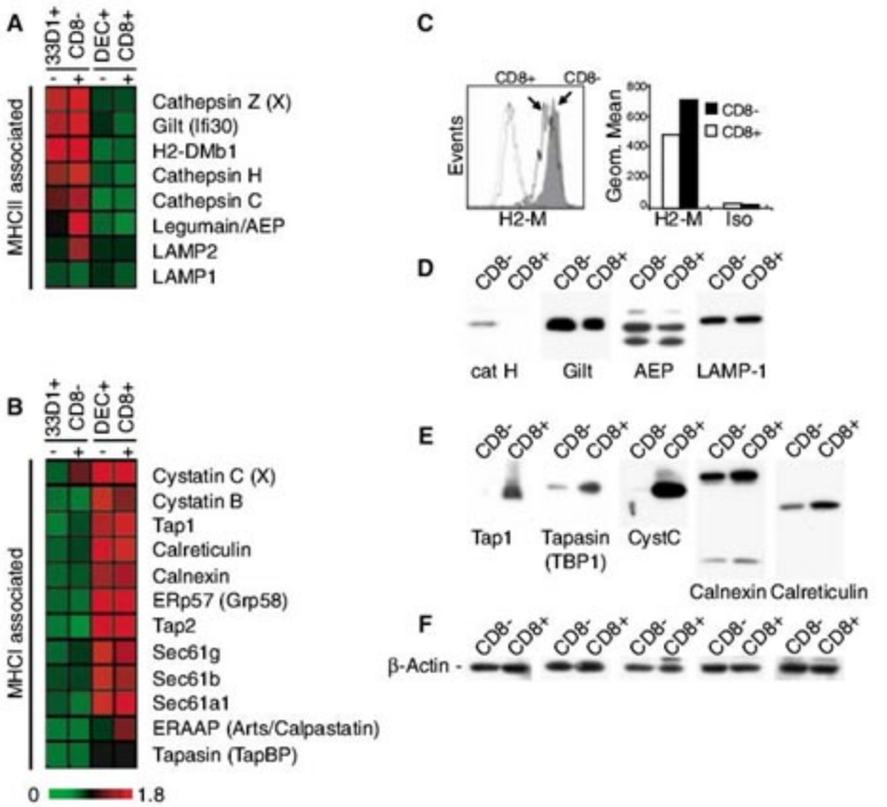


Fig. 4. Distinct expression pattern of MHC class I– and MHC class II–associated molecules. Affymetrix gene array analysis showing relative amounts of mRNAs associated with the MHC class II (**A**) and MHC class I (**B**) processing pathways expressed by $CD8^+DEC205^+$ and $CD8^-33D1^+$ DCs purified from WT- (–) and Flt3L-melanoma– (+) injected mice. Each bar represents the mean of three individual gene arrays prepared from distinct mRNA samples. (**C**) Intracellular FACS analysis of H2-DMb1/H2-DMa heterodimer in the $CD8^+DEC205^+$ and $CD8^-33D1^+$ DCs. (**D**) Western blots for cathepsin H, Gilt, and AEP on extracts of purified $CD8^+DEC205^+$ and $CD8^-33D1^+$ spleen DCs. (**E**) Western blots for Tap-1, tapasin, cystatin C, calnexin, and calreticulin on extracts of purified $CD8^+DEC205^+$ and $CD8^-33D1^+$ spleen DCs. (D) Lysosomal marker LAMP-1 (lysosomal-associated membrane protein 1) and (**F**) β-actin are shown as loading controls.

Efficient antigen presentation by DCs requires regulated lysosomal protein degradation (*21*, *22*). However, the requirements for presentation on MHCII and cross-presentation on MHCI differ in that MHCII processing occurs inside endosomes, whereas cross-presentation on MHCI necessitates antigen escape from the endosome into the cytoplasm to gain access to the proteasome and TAP transporters (*19*, *20*, *23*–*25*). Elegant in vitro experiments with cultured DCs show that during DC development, antigen presentation is regulated through control of lysosomal processing and MHCII cell surface transport (*21*, *22*, *26*–*28*). Cultured immature DCs capture antigen but only process and present it on MHCII after exposure to inflammatory stimuli or TLR ligation (*22*). This unique ability to sequester antigens may be important for their preservation during DC transit from sites of inflammation to lymphoid organs and might facilitate the escape of antigen from endosomes to the cytoplasm or endoplasmic reticulum for cross-presentation (*21*). However, DCs that fail to degrade antigen might also be suboptimal producers of MHCII-p. Our experiments show that in the intact host, this problem is resolved by producing a subset of DCs specialized for maximizing MHCII presentation. Although $CD8^+DEC205^+$ DCs can initiate immune responses by presenting on MHCII, $CD8^-33D1^+$ DCs excel in producing MHCII-p. This specialization may have important implications for understanding the initiation of T cell responses in vivo and for rational vaccine design.

References and Notes

1. Y. J. Liu, *Cell* **106**, 259 (2001).
2. K. Shortman, Y. J. Liu, *Nat. Rev. Immunol.* **2**, 151 (2002).
3. R. M. Steinman, D. Hawiger, M. C. Nussenzweig, *Annu. Rev. Immunol.* **21**, 685 (2003).
4. D. Vremec *et al.*, *J. Exp. Med.* **176**, 47 (1992).
5. M. D. Witmer, R. M. Steinman, *Am. J. Anat.* **170**, 465 (1984).
6. R. S. Allan *et al.*, *Science* **301**, 1925 (2003).
7. J. M. den Haan, S. M. Lehar, M. J. Bevan, *J. Exp. Med.* **192**, 1685 (2000).
8. T. Iyoda *et al.*, *J. Exp. Med.* **195**, 1289 (2002).
9. J. M. den Haan, M. J. Bevan, *J. Exp. Med.* **196**, 817 (2002).
10. J. L. Pooley, W. R. Heath, K. Shortman, *J. Immunol.* **166**, 5327 (2001).
11. C. Scheinecker, R. McHugh, E. M. Shevach, R. N. Germain, *J. Exp. Med.* **196**, 1079 (2002).
12. P. Schnorrer *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 10729 (2006).
13. Materials and methods are available as supporting material on *Science* Online.
14. D. Hawiger *et al.*, *J. Exp. Med.* **194**, 769 (2001).
15. M. C. Nussenzweig, R. M. Steinman, M. D. Witmer, B. Gutchinov, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 161 (1982).
16. L. Bonifaz *et al.*, *J. Exp. Med.* **196**, 1627 (2002).
17. G. Dadaglio, C. A. Nelson, M. B. Deck, S. J. Petzold, E. R. Unanue, *Immunity* **6**, 727 (1997).
18. M. Guo *et al.*, *Hum. Immunol.* **61**, 729 (2000).
19. P. Bryant, H. Ploegh, *Curr. Opin. Immunol.* **16**, 96 (2004).
20. P. Cresswell, N. Bangia, T. Dick, G. Diedrich, *Immunol. Rev.* **172**, 21 (1999).
21. L. Delamarre, M. Pack, H. Chang, I. Mellman, E. S. Trombetta, *Science* **307**, 1630 (2005).
22. E. S. Trombetta, I. Mellman, *Annu. Rev. Immunol.* **23**, 975 (2005).
23. R. N. Germain, *Cell* **76**, 287 (1994).
24. N. Shastri, S. Cardinaud, S. R. Schwab, T. Serwold, J. Kunisawa, *Immunol. Rev.* **207**, 31 (2005).
25. J. W. Yewdell, C. C. Norbury, J. R. Bennink, *Adv. Immunol.* **73**, 1 (1999).
26. K. Inaba *et al.*, *J. Exp. Med.* **191**, 927 (2000).
27. P. Pierre, I. Mellman, *Cell* **93**, 1135 (1998).
28. S. J. Turley *et al.*, *Science* **288**, 522 (2000).

29. We thank K. Velinzon for cell sorting; K.-H. Yao for technical assistance; H.-K. Lee, S. H. Park, and S. Y. Joe for help generating CD11c-hDEC205 mice; A. Flores-Langarica, S. Boscardin, A. Gazumyan, E. Besmer, and F. Nimmerjahn for helpful discussion; the MSKCC Monoclonal Antibody Core Facility for labeling of antibodies; and MSKCC Genomics Core Laboratory for performing the microarrays. Supported by grants from the NIH (to C.G.P., R.M.S., and M.C.N.). D.D. is a fellow of the Deutsche Forschungsgemeinschaft (DU 548/1-1), and V.R.B. was supported by the German National Academic Foundation. M.C.N. is an Investigator with the Howard Hughes Medical Institute. M.C.N. and R.M.S. are on the scientific advisory board of Celldex, a startup company interested in targeting dendritic cells.

Supporting Online Material

www.sciencemag.org/cgi/content/full/315/5808/107/DC1
Materials and Methods
Figs. S1 to S11
References

10 October 2006; accepted 15 November 2006
10.1126/science.1136080

Differential Transmission of Actin Motion Within Focal Adhesions

Ke Hu,* Lin Ji,* Kathryn T. Applegate, Gaudenz Danuser,† Clare M. Waterman-Storer†

Cell migration requires the transmission of motion generated in the actin cytoskeleton to the extracellular environment through a complex assembly of proteins in focal adhesions. We developed correlational fluorescent speckle microscopy to measure the coupling of focal-adhesion proteins to actin filaments. Different classes of focal-adhesion structural and regulatory molecules exhibited varying degrees of correlated motions with actin filaments, indicating hierarchical transmission of actin motion through focal adhesions. Interactions between vinculin, talin, and actin filaments appear to constitute a slippage interface between the cytoskeleton and integrins, generating a molecular clutch that is regulated during the morphodynamic transitions of cell migration.

Directed cell migration involves spatiotemporal orchestration of protrusion at the leading cell edge, adhesion of the protrusion to the extracellular matrix (ECM), pulling against the adhesions to translocate the cell body, and weakening of the adhesion at the cell rear for advancement (*1*). In this process, actin filaments (F-actin) must couple to the ECM through the plasma membrane (*1*–*3*) via focal adhesions (FAs) to translate actin polymerization and/or actin-myosin contraction into cell motion. FAs are complexes of >100 different proteins linking F-actin to clustered transmembrane integrin ECM receptors (*2*, *4*). Regulating the attachment between F-actin and integrins via proteins within FAs is thought to be critical for controlling the spatiotemporal variability of protrusion and traction (*5*) and the ability of cells to respond to mechanical cues.

It is well established that F-actin and FAs are coupled to each other. Many FA proteins bind directly or indirectly to F-actin (*6*–*8*) and/or integrins (*9*–*13*). Contractile actomyosin bundles are often rooted in FAs (*2*, *4*), and perturbations of actomyosin cause changes in FAs and vice versa (*2*). Although the importance of spatiotemporal coordination between FAs and F-actin in cell migration is well appreciated (*2*, *14*, *15*), it is not known which FA molecules interact with F-actin in living cells, and the dynamics of molecules within these two assemblies have never been analyzed simultaneously. Predicting how FA proteins behave in vivo by biochemical data alone is impossible because of the complexity of their interactions (*4*).

To study the dynamic interactions between F-actin and FAs, we combined total internal reflection fluorescence microscopy (TIRFM) and fluorescent speckle microscopy (FSM). TIRFM optimizes image contrast at the ventral cell–ECM/coverslip interface where cortical F-actin integrates with FAs. FSM marks macromolecular assemblies with fluorophore clusters called speckles (fig. S1). Computational tracking of speckle motion allows mapping of protein dynamics with submicron resolution (*16*, *17*). We studied PtK1 cells migrating on coverslips, on which they organized a fibronectin-containing ECM (fig. S2).

To determine the spatial relations between FAs and F-actin flow, we captured image pairs using TIRFM of green fluorescent protein

Department of Cell Biology, The Scripps Research Institute, La Jolla, CA 92037, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: waterman@scripps.edu (C.M.W.-S.); gdanuser@scripps.edu (G.D.)

(GFP)–vinculin and TIR-FSM of X-rhodamine actin (Fig. 1). As seen previously (*16*), the narrow meshwork of F-actin in the lamellipodium underwent rapid retrograde flow from the leading edge toward the cell center. Proximal to this, F-actin retrograde flow in the lamella was slower. An overlay of FAs onto the F-actin flow map revealed that the negative flow speed gradient at the lamellipodium/lamella junction corresponded to the distal boundaries of FAs (Fig. 1, C and D). Thus, FAs may locally dampen flow by engaging F-actin to the ECM. The global slowing of F-actin flow in the lamella is probably due to the uniform distribution of FAs and the small inter-FA spacing in this cell region. Despite the slowing, F-actin retrograde flow within FAs was substantial and coherent (Fig. 1C) (*18*). Thus, if an interaction between FA proteins and F-actin occurs in living cells, movement of proteins within FAs is likely.

To analyze the motion of proteins within FAs, we performed TIR-FSM on cells expressing three classes of GFP-conjugated FA proteins: first, a fibronectin-binding integrin (GFP–α_v integrin coexpressed with untagged β_3 integrin); second, FA proteins capable of binding directly to F-actin [α-actinin (*6*), vinculin (*8*, *19*), and talin (*6*, *20*)]; and third, FA "core" proteins that do not bind F-actin or the ECM directly but include structural and signaling molecules [paxillin (*9*, *21*), zyxin (*12*), and focal-adhesion kinase (FAK) (*13*)]. We focused on FAs in the ~100-nm-thick leading-edge lamella, where F-actin forms transverse bundles, isotropic networks, and stress fibers (*22*), all of which are within the evanescent excitation field. FA speckles were tracked within segmented FA regions, and average speckle speeds and coherencies were computed (*18*). Monte Carlo simulations showed that random binding and dissociation of fluorescent molecules to and from an immobile FA produced speckle velocities <0.05 μm/min and coherencies <0.4 on a scale of 0 to 1 (fig. S7), similar to measurements of X-rhodamine–actin speckles imaged in a fixed cell (*18*) and thus defining the detection limit of our measurements.

We found highly diverse behaviors of the seven GFP–FA proteins within FAs (Fig. 2 and tables S1 and S2). Speckles consisting of FA proteins with no known F-actin–binding activity moved slowly and mostly incoherently (Fig. 2, movies S1 to S3, and tables S1 and S2). Of these, GFP–$\alpha_v\beta_3$ integrin was the slowest and the most incoherent (Fig. 2C and movie S1), probably due to its immobilization by binding to the ECM on the coverslip. FA core proteins were slightly more motile than $\alpha_v\beta_3$ integrin, with a retrograde directional bias in their movement (Fig. 2, D to F). GFP-zyxin and GFP-paxillin moved somewhat more coherently than GFP–$\alpha_v\beta_3$ integrin, whereas GFP-FAK speckles moved faster than GFP–$\alpha_v\beta_3$ integrin (Fig. 2, D to F; movie S2; and tables S1 and S2).

In contrast, all three GFP-tagged FA actin–binding proteins moved coherently within the FAs. The dynamics of FA actin–binding proteins were significantly different from those of both the

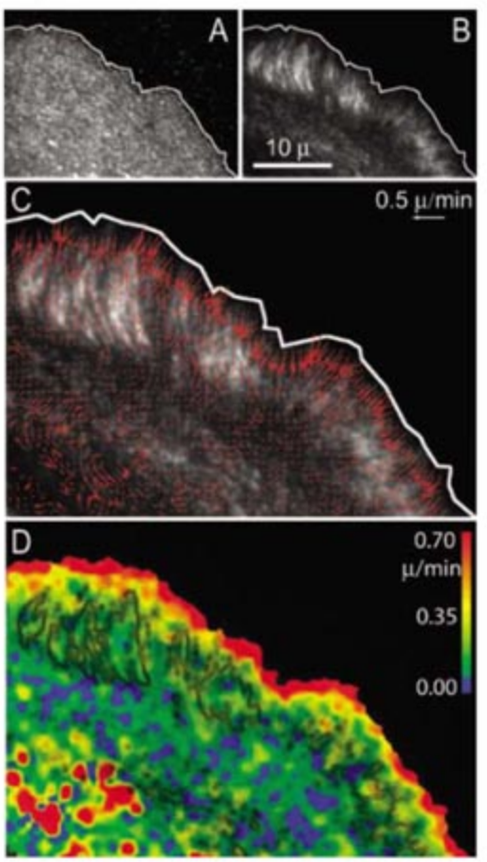


Fig. 1. F-actin motion relative to FAs at the leading edge of a migrating PtK1 epithelial cell. (**A**) TIR-FSM image of X-rhodamine actin and (**B**) TIRF image of GFP-vinculin. (**C**) Velocity vectors of F-actin speckle motion averaged over 100 s are overlaid on the TIRF image of GFP-vinculin. (**D**) Color-coded map of F-actin speed, with FAs outlined in gray.

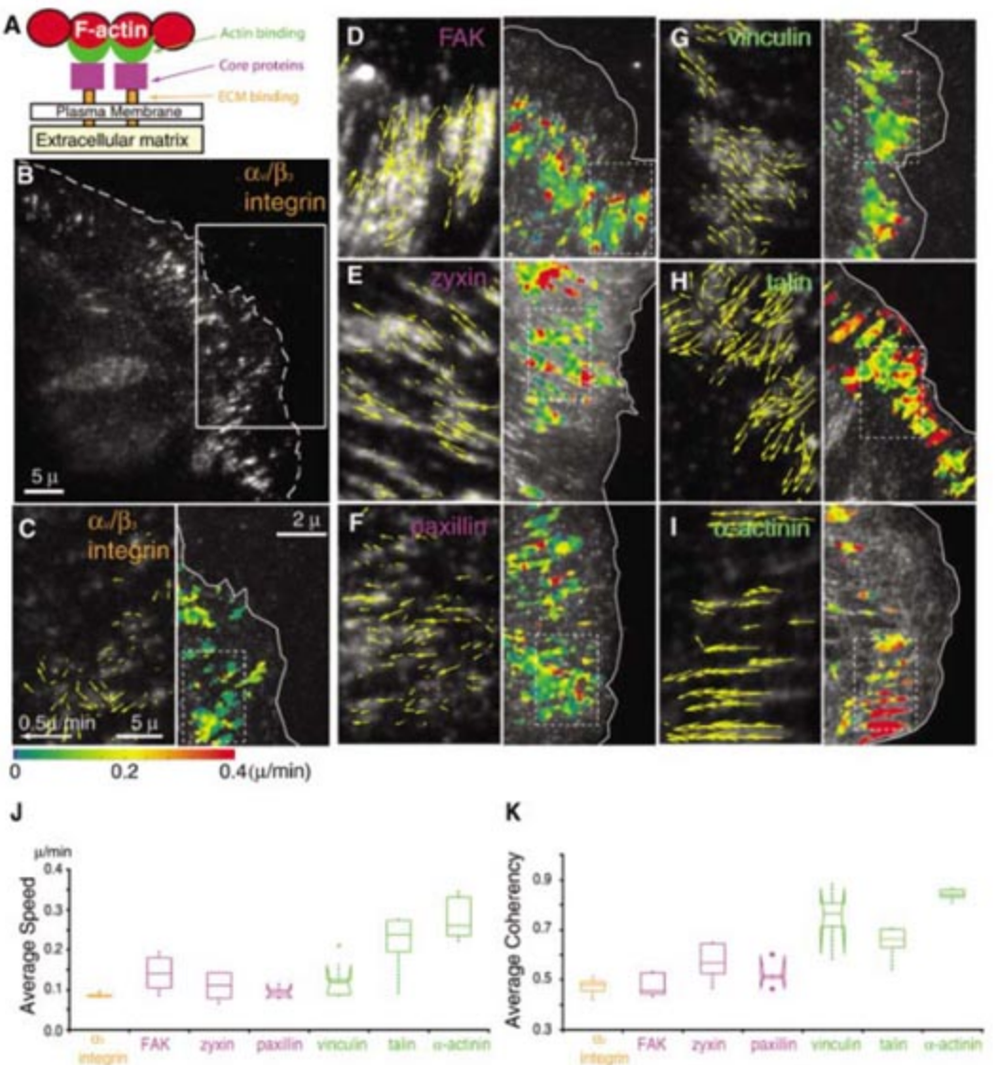


Fig. 2. The motions of different proteins within FAs are diverse. (**A**) Classes of FA molecules analyzed. FA-actin–binding proteins, green; FA core proteins, purple; integrins, orange. (**B**) TIR-FSM image of a cell coexpressing GFP–α_v integrin and untagged β_3 integrin. White frame, region shown in (C), on the right. (**C** to **I**) Velocity vectors (left) and speed maps (right, boxed area shown at left) of seven GFP-tagged FA proteins analyzed by TIR-FSM. (**J** and **K**) Average speed and velocity coherency of FA molecules [also see tables S1 and S2 (*18*) and fig. S3A].

integrin and core proteins and also from each other (Fig. 2, G to I; movie S3; and tables S1 and S2). GFP–α-actinin speckles moved fastest and most coherently. GFP-vinculin speckles moved slowly but highly coherently, whereas GFP-talin speckle motion was the least coherent of all three, but it was significantly faster than vinculin. Speed maps revealed that speckle speeds could vary within individual FAs and between adjacent FAs (Fig. 2, C to I). None of these FA proteins possess motor activity or interact with motor proteins, which suggests that their motion is influenced by interactions with other dynamic cell components, such as F-actin.

To determine whether the motion of proteins within FAs was related to F-actin flow, we developed correlational FSM to quantify the degree of motion correlation between GFP-FA and X-rhodamine–F-actin speckles (*18*). As verified by Monte Carlo simulations, a high degree of speckle motion correlation indicates concerted movement of molecules as part of the same macromolecular ensemble, mediated by direct or indirect interactions. We tracked speckle motion within FAs in both channels and interpolated speckle velocities onto a common 0.45-μm–by–0.45-μm grid to allow comparison of pairs of FA and F-actin flow vectors (fig. S3).

To determine the dependency between the velocities of FA protein and F-actin speckles, we performed linear regression of scatter plots of FA versus F-actin speckle velocities averaged within individual FAs (Fig. 3A and fig. S5A). This revealed velocity correlations between F-actin and the FA actin–binding proteins vinculin, α-actinin, and talin. The relatively low correlation coefficient for talin indicated a high variability in the talin–F-actin interaction.

To estimate the extent of molecular coupling between F-actin and FA molecules, we computed two parameters: direction coupling score (DCS = cos θ, where θ is the angle between paired FA and F-actin vectors), to measure the directional similarities between FA and F-actin speckle motions, and velocity magnitude coupling score (VMCS) (Fig. 3B and fig. S3), to measure relative FA speckle motion along the local F-actin flow axis, thus accounting for both direction and speed. For identical speckle flow fields, both DCS and VMCS are equal to 1 (see Fig. 3, G to H; tables S3 to S5; and fig. S4 for analysis of GFP- and Alexa 568–actin in the same cell).

Analysis of FAs in the lamella revealed that couplings between different FA molecules and F-actin were highly diverse (tables S4 and S5). FA core proteins and $\alpha_v\beta_3$ integrin both exhibited a DCS and a VMCS much less than 1 (Fig. 3, fig. S5, movies S5 to S8, and tables S4 and S5). Therefore, a substantial portion of the motion of

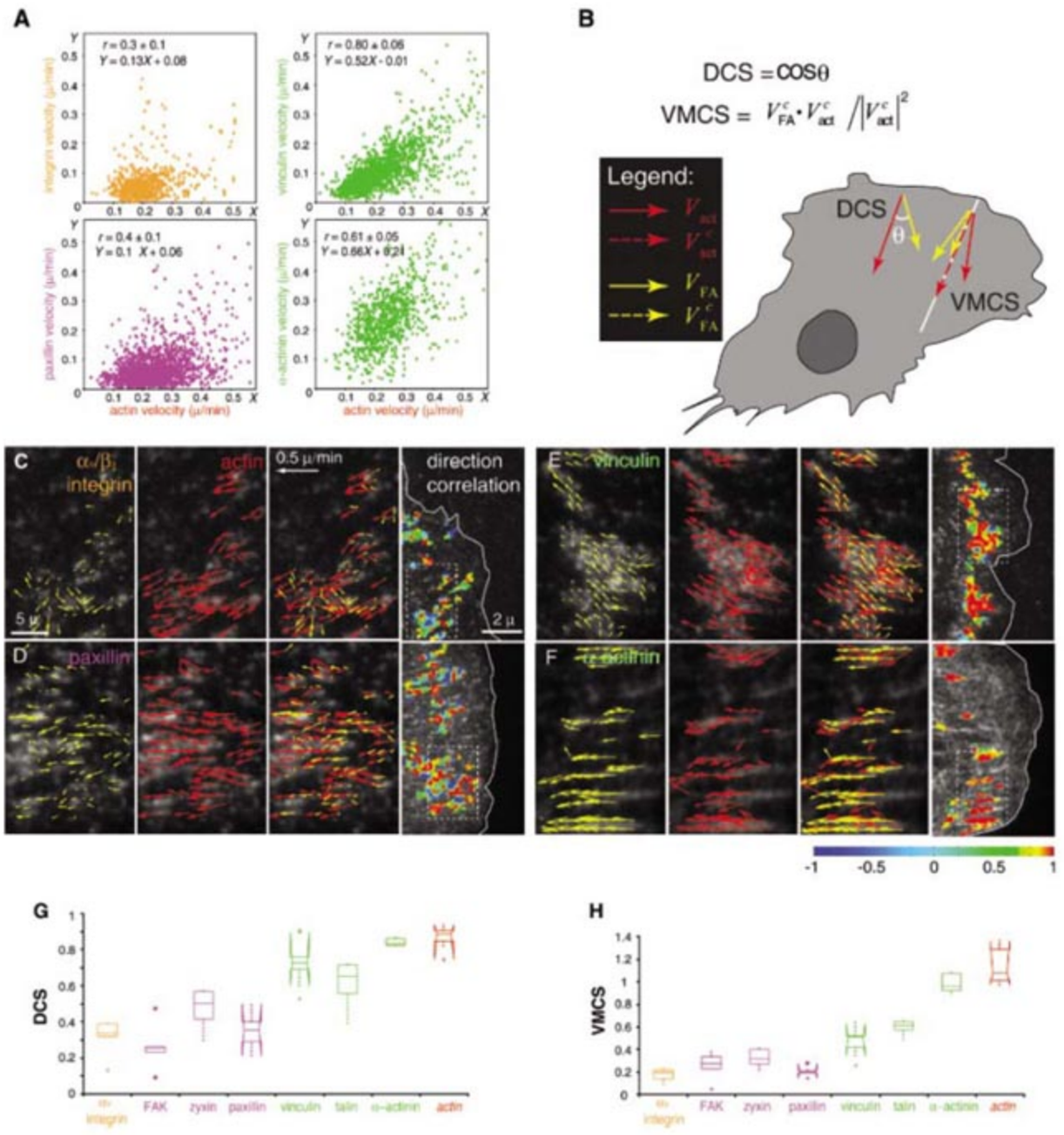


Fig. 3. Correlational FSM reveals that proteins within FAs are differentially coupled to F-actin motion. (**A**) FA versus F-actin speckle motions. Each point represents the average FA and F-actin speckle velocities within one FA at one time step. Correlation coefficients (*r*) and two times the standard deviation are indicated (bootstrap regression, 200 trials). *Y*, FA protein velocity; *X*, F-actin velocity. (**B**) Definitions of DCS and VMCS [also see (*18*) and fig. S3]. V_{FA} and V_{act} are the actual velocity measurements, and V^c_{FA} and V^c_{act} are the coupled components of the flow vectors. (**C** to **F**) Correlational FSM analysis of GFP–$\alpha_v\beta_3$ integrin (C), paxillin-GFP (D), vinculin-GFP (E), and α-actinin–GFP (F). From left to right in each panel are FA speckle velocity, F-actin speckle velocity, an overlay of the two velocities, and a color-coded DCS map. (**G** and **H**) Average DCS (G) and VMCS (H) between FA molecules and F-actin (averages from several cells, tables S4 and S5). Red, scores for a cell containing GFP-actin and Alexa 568–actin (fig. S4).

FA core proteins and $\alpha_v\beta_3$ integrin was not related to F-actin flow but was probably caused by interactions with other binding partners within FAs that immobilize them or deviate their motion from the F-actin flow axis. Alternatively, binding and dissociation of FA proteins to and from FAs could also generate random minor speckle displacements (figs. S6 and S7).

The motions of FA actin–binding proteins within FAs all showed significantly greater coupling to F-actin motion than did core proteins and $\alpha_v\beta_3$ integrin, although they were clearly different from each other (movies S9 to S11 and tables S4 and S5). α-Actinin displayed the highest coupling to F-actin motion (movie S11). This is expected because α-actinin mimics the kinematics of F-actin throughout the cell, indicating its tight association with F-actin, irrespective of localization (*23*, *24*). Both vinculin and talin were significantly, yet partially, coupled to F-actin motion, indicating partial transmission of F-actin motion to these proteins within FAs (tables S4 and S5).

Mapping local DCS revealed heterogeneity in coupling between F-actin and FA proteins within individual FAs and between adjacent FAs (Fig. 3, C to F). To see whether this heterogeneity was related to whole-FA dynamics or cell migration behavior, we performed correlational FSM for vinculin and F-actin (Fig. 4A and movie S12) at the leading edge of a cell where one area protruded and an adjacent area retracted (Fig. 4A). The protrusive area contained a FA that remained stationary, whereas the FA in the retracting area slid rearward and later stabilized (Fig. 4A). In the stationary FA, the speeds of F-actin and vinculin speckle flow and the DCS and VMCS between F-actin and vinculin remained relatively constant, with only small fluctuations over time (Fig. 4B). In contrast, in the sliding FA, the speeds of actin and vinculin and their VMCS and DCS increased before FA sliding (Fig. 4C). The coupling between vinculin and F-actin peaked during FA sliding and decreased before FA stabilization. Thus, dissociation of vinculin from a less mobile FA component and stable vinculin–F-actin binding may initiate FA disengagement from the ECM, whereas partial coupling between vinculin and F-actin may be necessary for establishing and/or maintaining the engagement between the FA and the ECM.

Our direct analysis of the dynamic interactions between FA components and F-actin in living cells reveals that the efficiency of motion transmission from F-actin to FA proteins within FAs decreased from actin-binding proteins to FA core proteins to integrin, defining a hierarchical slippage clutch. This is likely to be the result of differential transmission of F-actin–based force through a network of transient protein-protein interactions in FAs. Partial coupling of talin and vinculin to F-actin motion could represent these molecules spending part of

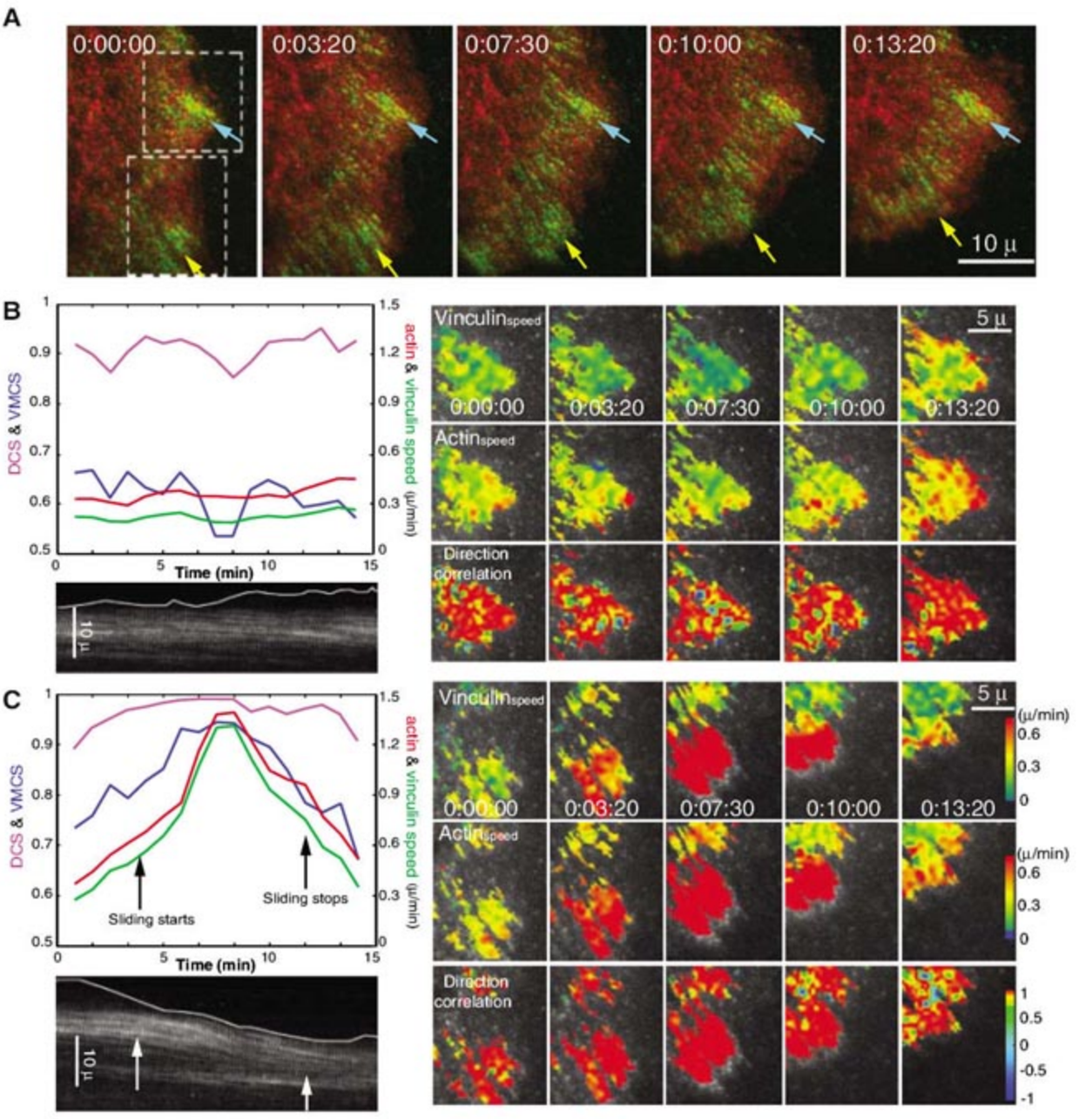


Fig. 4. Vinculin–F-actin coupling is time-modulated during the retraction of a FA. (**A**) Images of GFP-vinculin (green) and X-rhodamine–actin (red). Blue arrow, stable FA in a protrusive cell region; yellow arrow, sliding FA in a retracting cell region; white frames, regions of interest analyzed by FSM in (B) and (C). (**B** and **C**) Temporal variation of F-actin and vinculin speckle speeds, DCS, and VMCS within a stable (B) and a sliding (C) FA. Top left panels show graphs of average speeds of F-actin (red) and vinculin (green) speckles, vinculin-actin VMCS (blue), and vinculin-actin DCS (pink). Bottom left panels show kymographs of GFP-vinculin taken along the axes of arrows in (A). The position of the cell edge in white shows that the FA remains stationary in (B), whereas in (C) the FA initiates sliding at ~4 min (left arrow) and stops at ~12 min (right arrow). Right panels show maps of vinculin and actin speckle speeds and DCS. During retraction and FA sliding, vinculin alters its binding to F-actin. Time is given in hour:min:sec.

their time bound to moving F-actin and part of their time bound to a less mobile FA component, thus identifying these proteins as a site of slippage in the F-actin/FA interface. Alternatively, differential coupling of FA proteins to transverse actin bundles and stress fibers in the lamella could contribute to the observed effect. However, given the local slowing of F-actin flow in the FA in the lamella and the biophysical evidence implicating talin and vinculin in force transmission in the FA (*20*, *25*, *26*), we suspect that these proteins form transient linkages across the slippage interface, resulting in force-transducing slip-stick friction between F-actin and the ECM. The degree of molecular motion transmission through the FA was regulated, and it was correlated with protrusion and retraction events during cell migration. Therefore, FA internal molecular kinematics may be a key element in the integrin-mediated translation of intracellular biochemistry into cellular mechanics during cell and tissue morphogenesis, or in the reception of extracellular mechanical signals to mediate sensory perception, tissue maintenance, and differentiation (*27*).

References and Notes

1. D. A. Lauffenburger, A. F. Horwitz, *Cell* **84**, 359 (1996).
2. K. Burridge, M. Chrzanowska-Wodnicka, *Annu. Rev. Cell Dev. Biol.* **12**, 463 (1996).
3. C. H. Lin, P. Forscher, *Neuron* **14**, 763 (1995).
4. B. Geiger, A. Bershadsky, R. Pankov, K. M. Yamada, *Nat. Rev. Mol. Cell Biol.* **2**, 793 (2001).
5. R. J. Pelham Jr., Y. Wang, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 13661 (1997).
6. K. Maruyama, S. Ebashi, *J. Biochem. (Tokyo)* **58**, 13 (1965).
7. M. Muguruma, S. Matsumura, T. Fukazawa, *Biochem. Biophys. Res. Commun.* **171**, 1217 (1990).
8. R. P. Johnson, S. W. Craig, *Nature* **373**, 261 (1995).
9. T. Tanaka, R. Yamaguchi, H. Sabe, K. Sekiguchi, J. M. Healy, *FEBS Lett.* **399**, 53 (1996).
10. D. A. Calderwood *et al.*, *J. Biol. Chem.* **274**, 28071 (1999).
11. K. Burridge, P. Mangeat, *Nature* **308**, 744 (1984).
12. M. C. Beckerle, *Bioessays* **19**, 949 (1997).
13. D. D. Schlaepfer, T. Hunter, *Cell Struct. Funct.* **21**, 445 (1996).
14. K. M. Yamada, S. Miyamoto, *Curr. Opin. Cell Biol.* **7**, 681 (1995).
15. M. Chrzanowska-Wodnicka, K. Burridge, *J. Cell Biol.* **133**, 1403 (1996).
16. A. Ponti, M. Machacek, S. L. Gupton, C. M. Waterman-Storer, G. Danuser, *Science* **305**, 1782 (2004).
17. L. Ji, G. Danuser, *J. Microsc.* **220**, 150 (2005).
18. Materials and methods are available as supporting material on *Science* Online.
19. B. Geiger, *Cell* **18**, 193 (1979).
20. G. Jiang, G. Giannone, D. R. Critchley, E. Fukumoto, M. P. Sheetz, *Nature* **424**, 334 (2003).
21. C. E. Turner, J. R. Glenney Jr., K. Burridge, *J. Cell Biol.* **111**, 1059 (1990).
22. S. L. Gupton *et al.*, *J. Cell Biol.* **168**, 619 (2005).
23. M. Edlund, M. A. Lotano, C. A. Otey, *Cell Motil. Cytoskelet.* **48**, 190 (2001).
24. G. Giannone *et al.*, *Cell* **116**, 431 (2004).
25. R. M. Ezzell, W. H. Goldmann, N. Wang, N. Parasharama, D. E. Ingber, *Exp. Cell Res.* **231**, 14 (1997).
26. G. Giannone, G. Jiang, D. H. Sutton, D. R. Critchley, M. P. Sheetz, *J. Cell Biol.* **163**, 409 (2003).
27. A. Katsumi, A. W. Orr, E. Tzima, M. A. Schwartz, *J. Biol. Chem.* **279**, 12001 (2004).

28. We thank C. Otey (University of North Carolina, Chapel Hill), A. Huttenlocher (University of Wisconsin, Madison), D. Schlaepfer (Scripps), A. F. Horwitz (University of Virginia, Charlottesville), I. Kaverina (Vanderbilt University), and M. Ginsberg (University of California, San Diego) for complementary DNAs. Supported by NIH grants GM67230 (C.M.W.-S. and G.D.) and U54GM64346 (G.D. and L.J.), American Heart Association Established Investigatorship and NIH Director's Pioneer Award (C.M.W.-S.), Leukemia and Lymphoma Society (K.H.), and NSF (K.T.A.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/315/5808/111/DC1
Materials and Methods
Figs. S1 to S8
Tables S1 to S5
References
Movies S1 to S14

13 September 2006; accepted 16 November 2006
10.1126/science.1135085

Live-Cell Imaging of Enzyme-Substrate Interaction Reveals Spatial Regulation of PTP1B

Ivan A. Yudushkin,[1]* Andreas Schleifenbaum,[1]* Ali Kinkhabwala,[1]* Benjamin G. Neel,[2] Carsten Schultz,[1] Philippe I. H. Bastiaens[1]†

Endoplasmic reticulum–localized protein-tyrosine phosphatase PTP1B terminates growth factor signal transduction by dephosphorylation of receptor tyrosine kinases (RTKs). But how PTP1B allows for RTK signaling in the cytoplasm is unclear. In order to test whether PTP1B activity is spatially regulated, we developed a method based on Förster resonant energy transfer for imaging enzyme-substrate (ES) intermediates in live cells. We observed the establishment of a steady-state ES gradient across the cell. This gradient exhibited robustness to cell-to-cell variability, growth factor activation, and RTK localization, which demonstrated spatial regulation of PTP1B activity. Such regulation may be important for generating distinct cellular environments that permit RTK signal transduction and that mediate its eventual termination.

Protein-tyrosine phosphorylation is widely used by eukaryotic cells to transduce signals, but the dynamic interplay between receptor tyrosine kinases (RTKs) and protein-tyrosine phosphatases (PTPs) remains poorly understood (*1*, *2*). The protein tyrosine phosphatase–1B (PTP1B) resides on the surface of the endoplasmic reticulum (ER) (*3*, *4*) and helps terminate signaling by multiple RTKs, including the epidermal growth factor receptor (EGFR) (*5*). Previous reports demonstrate that RTK signaling occurs at the plasma membrane and endosomes (*6*), and its termination occurs along the ER surface (*7*–*11*). Because PTP1B has much higher specific activity than typical RTKs in vitro (*12*, *13*), uniformly high PTP1B activity along the ER could prevent endosomal RTK signaling. To account for compartmentalized RTK signaling, we hypothesized that PTP1B might exist inside cells as spatially separated subpopulations with different kinetic properties.

To test this hypothesis, we developed an imaging approach based on Förster resonant energy transfer (FRET) to spatially resolve enzyme-substrate (ES) interactions and thereby to monitor enzyme activity in live cells (Fig. 1A) (*11*). We tagged PTP1B with a donor chromophore by fusion to a genetically encoded fluorescent protein, and conjugated the substrate, a synthetic phosphotyrosine-containing peptide, to an acceptor chromophore (Fig. 1B). For Michaelis-Menten kinetics, the steady-state fraction (α) of ES complex to total enzyme (E_0) is as follows:

$$\alpha = \mathrm{ES}/E_0 = S/(K_M + S) \qquad (1)$$

where S is the substrate concentration, and K_M is the Michaelis-Menten constant. The fraction α can be mapped across the cell by quantitatively imaging FRET with the use of fluorescence lifetime imaging microscopy (FLIM) (*14*–*16*).

We first tested whether formation of the ES intermediate could be detected by FRET in vitro. To stabilize the ordinarily transient ES intermediate and thereby to facilitate FRET detection, we used the purified enhanced green fluorescent protein (EGFP)–tagged catalytically impaired mutant of PTP1B that retains substrate-binding ability (residues 1 to 321, $PTP1B^{D181A}$, in which Ala^{181} was substituted for Asp) (*17*, *18*). Indeed, FRET in the ES complex was apparent, as in-

[1]European Molecular Biology Laboratory (EMBL), Meyerhofstrasse 1, D-69117 Heidelberg, Germany. [2]Cancer Biology Program, Division of Hematology-Oncology, Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: bastiaen@embl.de

dicated by increased acceptor (lissamine) emission, quenching of donor (EGFP) fluorescence (Fig. 1C), and decreased donor fluorescence lifetime (fig. S3). Consistent with the PTP catalytic mechanism, the thiol-reactive compounds iodoacetamide and sodium pervanadate (*19*), as well as the transition-state analog sodium vanadate, all attenuated FRET (Fig. 1D and fig. S3). By contrast, incubation of the peptide with the wild-type catalytic domain resulted in a transient FRET signal, which decayed with a velocity consistent with previously determined catalytic parameters [K_M of 2.2 μM, and the catalytic rate of the enzyme (k_{cat}) of 67 s^{-1}; fig. S4] (*13*, *20*).

Next, we analyzed interactions between PTP1B and the synthetic substrate in COS-7 cells, which endogenously express EGFR (*21*). We compared the EGFP fluorescence lifetime in cells ectopically expressing EGFP-PTP1B^{WT} with and without microinjected caged substrate (**1**-cP) (Fig. 2A). As expected, no FRET signal was observed before uncaging, because the cage on the phosphate moiety prevents ES interaction (*22*) (Fig. 2B, top). Ultraviolet (UV)–induced uncaging resulted in a significant drop of EGFP fluorescence lifetime, only in substrate-injected cells (Fig. 2B, middle), indicative of FRET. This interaction was specific for the phosphorylated peptide, as no FRET was detected when EGFR-deficient MCF-7 cells were loaded with the cognate nonphosphorylated peptide (see below). A transient, uniform distribution of ES complex was observed 15 s after uncaging, which evolved to a persistent gradient of ES complex by 1 min (Fig. 2C), with higher ES complex at the cell periphery. This slow equilibration time scale is inconsistent with the high k_{cat} measured in vitro (67 μM s^{-1} turnover for 1 μM of ES complex), which suggests that k_{cat} in vivo is <1% of that in vitro and/or only a small fraction (<1%) of PTP1B is fully active in vivo. The expected rapid equilibration time scale of <1 s was more rigorously confirmed by analysis of the time domain of a reaction-diffusion model (discussed below) that assumes the in vitro kinetic parameters.

The establishment of a stable ES complex concentration after stepwise uncaging of the substrate in COS-7 cells implied a steady-state maintained by a phosphorylation-dephosphorylation cycle (Fig. 2E). Notably, peptides containing the DADEYL phosphorylation motif can bind and be phosphorylated by the high basal kinase activity of the EGFR in the absence of epidermal growth factor (EGF) (*12*, *18*, *23*). To verify that ES complex was maintained at a steady state by a kinase-phosphatase reaction cycle, we tested whether nonphosphorylated substrate could be phosphorylated in cells. Loading a membrane-permeant, nonphosphorylated analog of the substrate (peptide **2**-OH) into MCF-7 cells ectopically coexpressing the citrine variant of yellow fluorescent protein fused to PTP1B^{D181A} (citrine-PTP1B^{D181A}) and cyan fluorescent protein fused to EGFR^{WT} (EGFR^{WT}-CFP) resulted in accumulation of substrate-PTP1B^{D181A} complex over

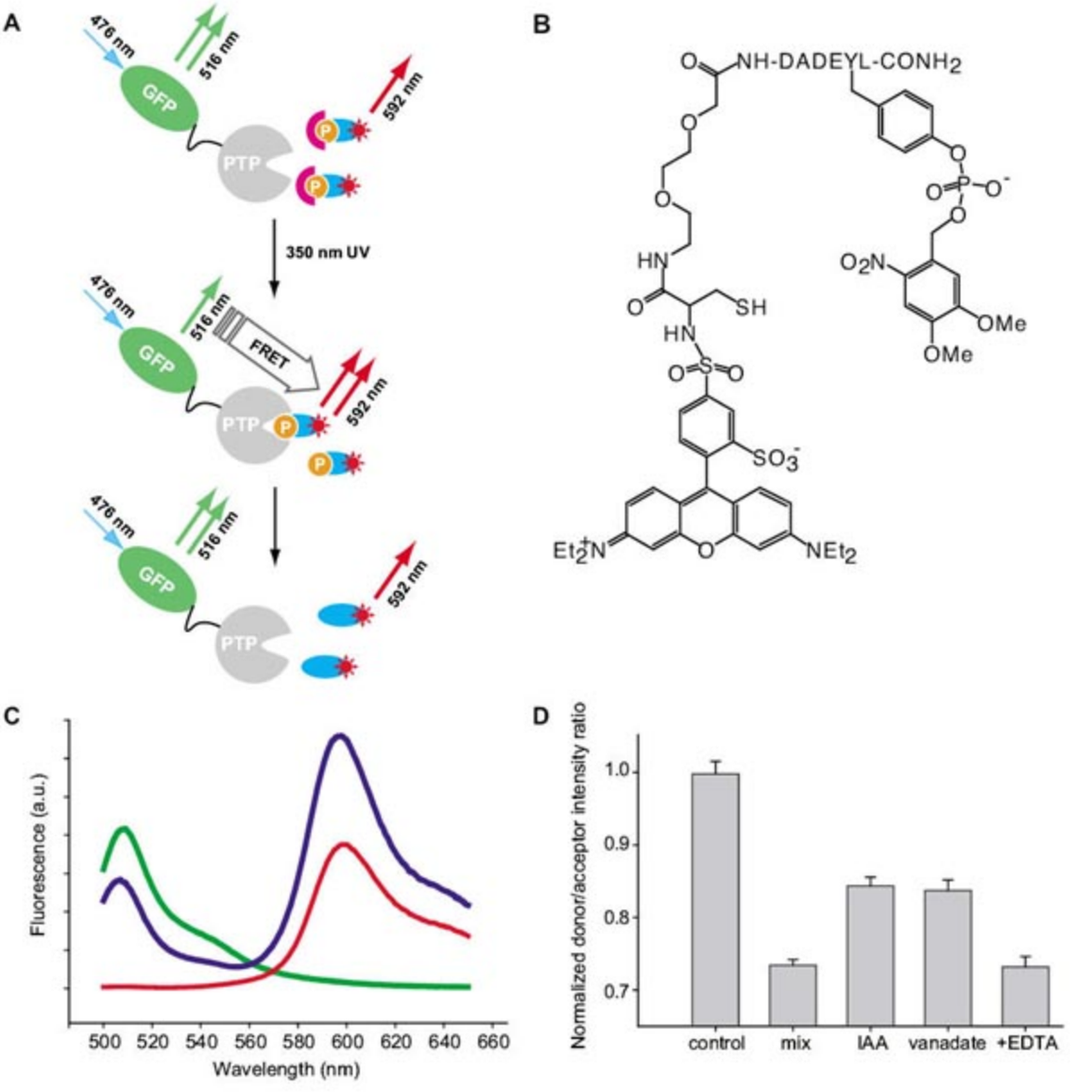


Fig. 1. Monitoring ES intermediate by FRET. (**A**) Assay format. (Top) A photolabile chemical protection group ("cage") on the phosphate moiety of the phosphotyrosine-containing synthetic peptide prevents binding to the active site of the PTP. (Middle) UV-induced photolysis of the cage induces substrate binding at the PTP active site and FRET, monitored by FLIM and/or emission intensity changes. (Bottom) After catalysis, the reaction product dissociates from the PTP, resulting in loss of FRET. (**B**) The synthetic hexapeptide substrate (LRh-Cys-AEEAc-DADEY^{NVOP}L-$CONH_2$, **2**-cP), corresponding to amino acids 988 to 993 of EGFR (phosphorylation at Y992), contains a photolabile 6-nitroveratryloxyphosphoryl (NVOP) group on the tyrosine residue and is coupled with FRET acceptor fluorophore lissamine rhodamine B (LRh) via an amino-ethyloxy-ethyloxy-acetyl (AEEAc) linker. An alternative peptide (**1**-cP) with an aminohexanoic (Ahx) acid linker and lacking the cysteine is not shown. For some live-cell experiments, **2**-cP was coupled to the cell internalization sequence from the third helix of the Antennapedia homeodomain (*26*). (**C**) Stabilization of the ES intermediate causes FRET between the EGFP-PTP1B(1–321)^{D181A} (green) and LRh-Ahx-DADEY^{P}L-$CONH_2$ (red) as seen by decreased emission of EGFP and increased fluorescence of lissamine on excitation at 476 nm (blue trace). (**D**) The thiol-reactive compound iodoacetamide (IAA) or the transition-state analog vanadate inhibits formation of the ES complex in vitro. EDTA chelates vanadate, reversing competitive inhibition of the substrate binding to EGFP-PTP1B(1–321)^{D181A}. Error bars show the SD of a typical experiment (n = 3).

time. This complex was not observed in the absence of $EGFR^{WT}$ or when the kinase-dead mutant $EGFR^{V741G}$ was transfected (Fig. 2F). Irreversible PTP inhibitor sodium pervanadate or the specific EGFR kinase inhibitor gefitinib (Iressa) also decreased FRET between citrine-tagged PTP1B and the substrate (figs. S7 and S8) in $EGFR^{WT}$-expressing MCF-7 cells, which confirmed the dynamics of the reaction cycle.

Because this reaction cycle establishes a steady-state concentration of the phosphorylated substrate, we could analyze the spatial distribution of ES complexes. Notably, the steady-state concentration of the ES complex was higher at the cell periphery than in the perinuclear region (Fig. 3, A to C). Using Eq. 1, a map of K_M/S for $PTP1B^{WT}$ was constructed (Fig. 3D), which revealed that the peripheral pool of PTP1B operated in a near-saturation regime (i.e., low K_M/S). An ES gradient dependent on K_M and S could reflect spatial regulation of PTP1B catalytic activity (by its effect on K_M). Alternatively, such a gradient could result from a gradient of phosphorylated peptide and/or differential competitive binding of PTP1B to endogenous substrates. The latter was ruled out by the uniform ES distribution observed in cells transfected with citrine-$PTP1B^{D181A}$ (Fig. 3, A to C), which binds the synthetic substrate reversibly and competitively (Fig. 1D and fig. S5).

Systems with spatially partitioned activities—e.g., a plasma membrane–bound kinase and a uniform ER-bound phosphatase (fig. S11)—are expected to generate steady-state substrate concentration gradients (*24*, *25*). In order to explore the contribution of a substrate gradient to the steady-state spatial distribution of an ES complex, we formulated a reaction-diffusion model of the proposed kinase-phosphatase reaction cycle (*11*). Our model, assuming the in vitro measured kinetic parameters, indeed generated a precipitous gradient (*25*) of the phosphorylated substrate close to the plasma membrane and also a gradient in ES complex (Fig. 4B). However, the maximal steady-state level of the ES complex predicted by the reaction-diffusion model was much lower than that observed experimentally (Fig. 3). Instead, the observed ES gradient could only be generated by the model if the specific activity (k_{cat}) of the phosphatase was <1% of that in vitro and/or a correspondingly small fraction (<1%) of PTP1B was fully active, consistent with our inference from the slow equilibration time scale shown in Fig. 2C. However, both solutions (the former shown in Fig. 4, A and B) display high sensitivity to relative levels of the kinase, the phosphatase, and the substrate (Fig. 4, A and B, and figs. S12 to S14) (*11*). This is incompatible with the persistence of the ES gradient to the observed variability of concentrations from cell to cell and implies an additional buffering mechanism in addition to reaction and diffusion.

Such a buffering mechanism could be provided by spatial regulation of the catalytic activity of PTP1B, which can well account for the observed robustness of the ES gradient over a large range of parameters (Fig. 4, C and D)

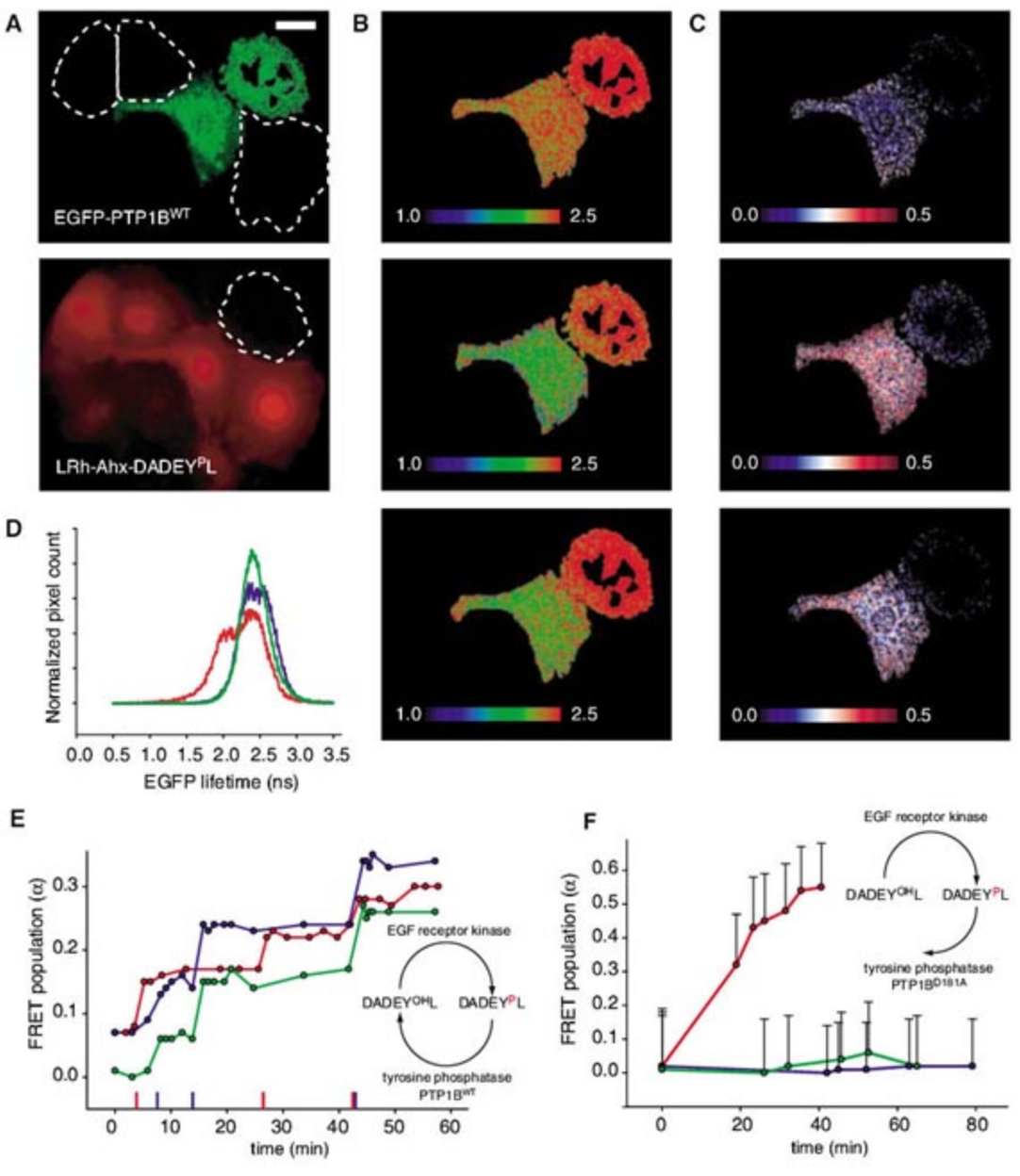


Fig. 2. Establishment of steady-state by kinase-phosphatase reaction cycle in live cells. (**A**) Reporter substrate LRh-Ahx-$DADEY^{NVOP}L$-$CONH_2$ (**1**-cP) (bottom) was microinjected into one of two COS-7 cells transiently expressing $EGFP\text{-}PTP1B^{WT}$ (top). Scale bar, 50 μm. (**B**) EGFP lifetime (τ_{phase}, 1.0 to 2.5 ns) images of $EGFP\text{-}PTP1B^{WT}$–expressing cells before (top), 15 s (middle), and 1 min (bottom) after 10-s irradiation with 360-nm UV. (**C**) Corresponding images of ES fraction (α) determined from global analysis of lifetime data (*16*). (**D**) Cumulative histogram of lifetime (τ_{phase}) values of the noninjected (green) and substrate-injected cell, before (blue) and 1 min after (red) UV-induced photolysis. (**E**) Stepwise 360-nm UV-induced uncaging results in higher steady-state level of the $EGFP\text{-}PTP1B^{WT}$-substrate complex in LRh-Ahx-$DADEY^{NVOP}L$-$CONH_2$–injected (peptide **1**-cP) COS-7 cells. Typical traces for three individual cells are shown. On the *x* axis, 10-s UV flashes are indicated as colored ticks. (**F**) Loading MCF-7 cells, cotransfected with the citrine-$PTP1B^{D181A}$ and CFP-tagged EGF receptor, with the membrane-permeable nonphosphorylated substrate LRh-C-AEEAc-$DADEY^{OH}L$-$CONH_2$ (**2**-OH) results in accumulation of the $PTP1B^{D181A}$-substrate complex over time (red), not seen with the citrine-$PTP1B^{D181A}$ alone (green) or in cells coexpressing the catalytically impaired EGFR receptor kinase mutant, $EGFR^{V741G}$-CFP (blue). Error bars show the SD of individual cell traces ($n = 5$).

(*11*). Here, the lower activity of the membrane-proximal PTP1B pool puts it into a near-saturation regime, whereas higher activity of the central pool buffers the system to perturbations over a larger range (Fig. 4, C and D, and figs. S15 to S17) (*11*). Spatial regulation also would explain the persistence of the ES gradient in cells with different shapes, as well as in cells with nonstimulated EGFR localized to endomembranes (fig. S8).

Persistence of the ES gradient was also observed after EGFR activation by EGF, which increases its kinase activity by up to a factor of 5 (*13*). At early times after stimulation, only a slight change in the gradient was observed (Fig. 3E), consistent with the spatial regulation model. Further increase in the steady-state level of ES complex in the cell interior observed at later times could be explained by the redistribution of activated EGFR after endocytosis.

Spatial regulation of PTP1B is likely important for its role in compartmentalized RTK signaling. In the absence of stimulus, the low activity of membrane-proximal PTP1B is still sufficient to counteract basal EGFR activity and, thereby, to maintain a low level of EGFR phosphorylation on endomembranes. With growth factor stimulation, the low activity of the plasma membrane–proximal PTP1B pool also may permit signaling from phosphorylated EGFR in endosomes. Transport of the ligand-activated EGFR to the perinuclear region (*8*, *9*), where PTP1B activity is higher, would then lead to signal termination. Further studies are required to elucidate the biochemical mechanism underlying the spatial regulation of PTP1B activity. Our method for imaging ES intermediates could be used to study the localization and regulation of potentially any enzymatic activity in live cells, providing quantitative information on enzyme catalysis with high specificity.

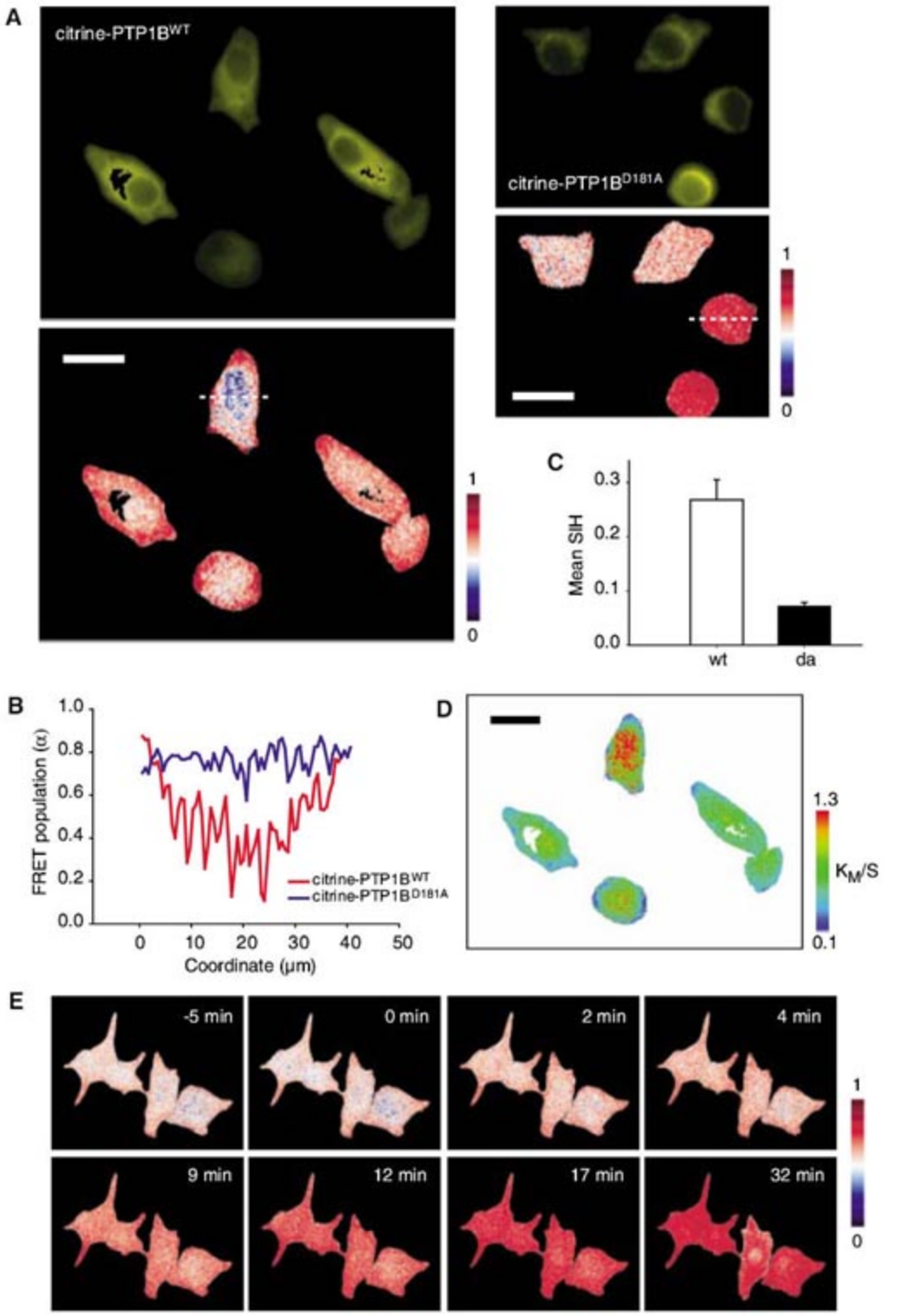


Fig. 3. Differential regulation of PTP1B activity in subcellular compartments. (**A**) Images of fraction (α) at steady state of ES complex over total citrine-PTP1B^{WT} (left) and citrine-PTP1B^{D181A} (right) in MCF-7 cells, coexpressing EGFR^{WT}-CFP and loaded with the membrane-permeable substrate LRh-C-AEEAc-DADEY^{P}L-CONH$_2$ (**2**-cP) linked to the penetratin peptide. (top) Fluorescence distribution of citrine-PTP1B. (**B**) Profile of fraction (α) across cells in (A) (dashed lines). (**C**) Statistical evaluation of the data in (A). The spatial inhomogeneity (SIH) score is calculated as described in (*27*). The error bars show SEM of SIH scored in 46 individual cells expressing citrine-PTP1B^{WT} (wt) and 47 cells expressing citrine-PTP1B^{D181A} (da) in at least four independent experiments. (**D**) Map of K_M/S for PTP1B^{WT}, calculated using Eq. 1. (**E**) Response of ES/E_0 after stimulation with 100 ng/ml EGF at 0 min in MCF-7 cells coexpressing citrine-PTP1B^{WT}/EGFR^{WT}-CFP and loaded with the membrane-permeable substrate **2**-cP linked to penetratin. Scale bars, 50 μm.

References and Notes

1. A. Ostman, F. D. Bohmer, *Trends Cell Biol.* **11**, 258 (2001).
2. P. Chiarugi, P. Cirri, *Trends Biochem. Sci.* **28**, 509 (2003).
3. J. V. Frangioni, P. H. Beahm, V. Shifrin, C. A. Jost, B. G. Neel, *Cell* **68**, 545 (1992).
4. T. A. Woodford-Thomas, J. D. Rhodes, J. E. Dixon, *J. Cell Biol.* **117**, 401 (1992).
5. A. Bourdeau, N. Dube, M. L. Tremblay, *Curr. Opin. Cell Biol.* **17**, 203 (2005).
6. M. Miaczynska, L. Pelkmans, M. Zerial, *Curr. Opin. Cell Biol.* **16**, 400 (2004).
7. M. Offterdinger, V. Georget, A. Girod, P. I. Bastiaens, *J. Biol. Chem.* **279**, 36972 (2004).
8. F. G. Haj, P. J. Verveer, A. Squire, B. G. Neel, P. I. Bastiaens, *Science* **295**, 1708 (2002).
9. Y. Romsicki, M. Reece, J. Y. Gauthier, E. Asante-Appiah, B. P. Kennedy, *J. Biol. Chem.* **279**, 12868 (2004).
10. N. Boute, S. Boubekeur, D. Lacasa, T. Issad, *EMBO Rep.* **4**, 313 (2003).
11. Materials and Methods are available as supporting material on *Science* Online.
12. Y. X. Fan, L. Wong, T. B. Deb, G. R. Johnson, *J. Biol. Chem.* **279**, 38143 (2004).
13. Z. Y. Zhang, D. Maclean, A. M. Thieme-Sefler, R. W. Roeske, J. E. Dixon, *Anal. Biochem.* **211**, 7 (1993).
14. P. J. Verveer, A. Squire, P. I. Bastiaens, in *Methods in Cellular Imaging*, A. Periasamy, Ed. (published for the American Physiological Society by Oxford Univ. Press, New York, 2001), pp. 273–294.
15. F. S. Wouters, P. J. Verveer, P. I. Bastiaens, *Trends Cell Biol.* **11**, 203 (2001).
16. P. J. Verveer, A. Squire, P. I. Bastiaens, *Biophys. J.* **78**, 2127 (2000).
17. A. J. Flint, T. Tiganis, D. Barford, N. K. Tonks, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1680 (1997).
18. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro;

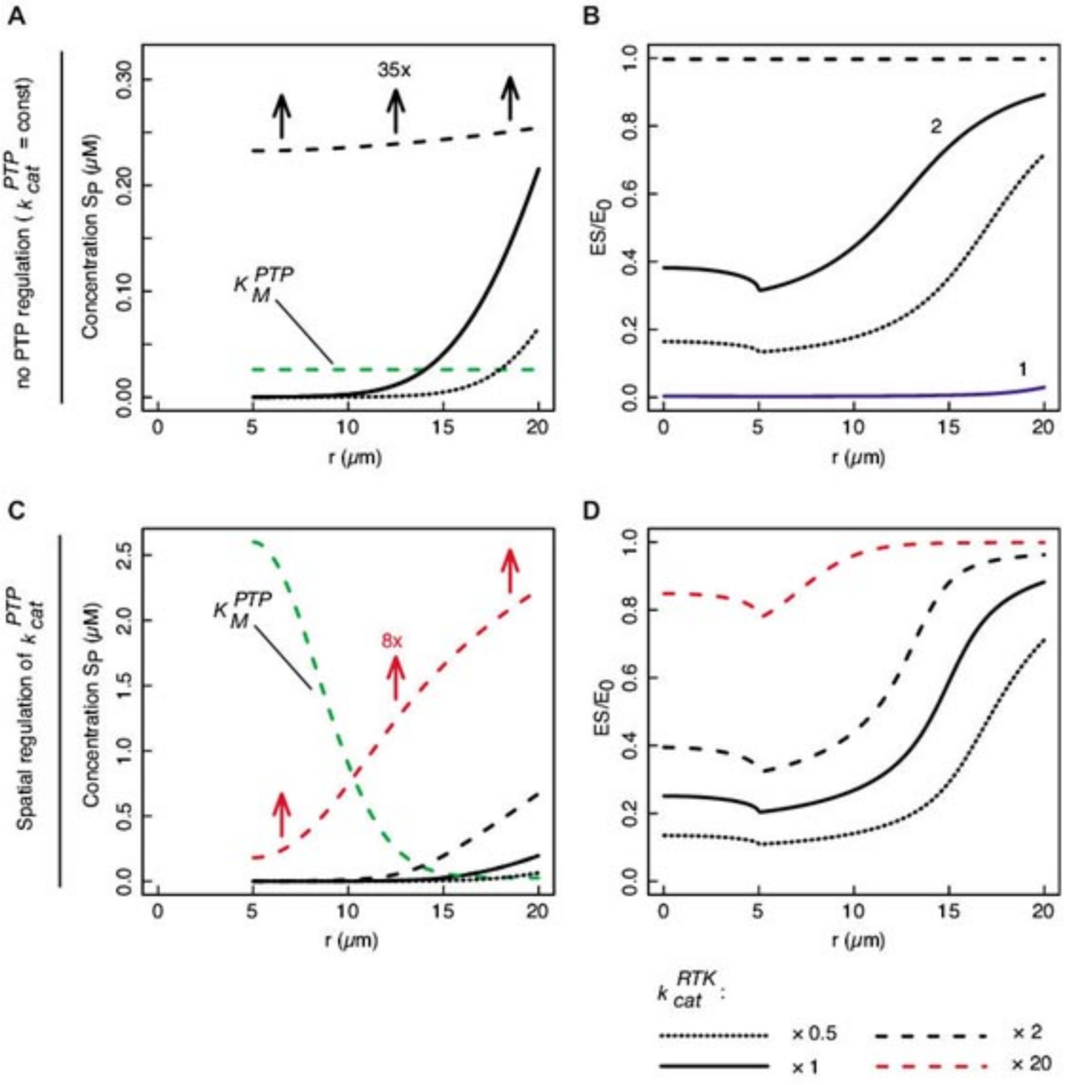


Fig. 4. Spherical reaction-diffusion models. (**A**) Phosphorylated substrate concentration (solid line) for a pure reaction-diffusion model using the values given in (*11*). The cytoplasm extends from the nucleus at 5 μm to the plasma membrane at 20 μm. All values are from observed or in vitro measurements, except for the phosphatase k_{cat}^{PTP} and K_M^{PTP} (dashed green line), which were both reduced by a factor of 100. The effect of a factor-of-2 increase (upper dashed curve, divided by 35) and decrease (lower dashed curve) in the kinase k_{cat}^{RTK} is also shown. (**B**) Two-dimensional projection of the ES/E_0 gradient for the corresponding curves in (A). Also shown is ES/E_0 assuming the in vitro values for all kinetic constants (curve 1). (**C**) Phosphorylated substrate concentration (solid line) for a reaction-diffusion model with additional catalytic regulation of k_{cat}^{PTP} and K_M^{PTP} (dashed green line) modeled as a Gaussian curve. At the plasma membrane, k_{cat}^{PTP} and K_M^{PTP} are lower than in vitro by 100; at the nuclear radius, k_{cat}^{PTP} and K_M^{PTP} assume their in vitro values. A factor-of-2 increase or decrease of the kinase k_{cat}^{RTK} is also shown (upper and lower dashed curves, respectively). Additionally, a factor-of-20 increase of the kinase k_{cat}^{RTK} is given (dashed red line, divided by 8). (**D**) Two-dimensional projection of ES/E_0 for the corresponding curves in (C).

Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; X, any amino acid; and Y, Tyr.
19. G. Huyer *et al.*, *J. Biol. Chem.* **272**, 843 (1997).
20. Z. Y. Zhang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 4446 (1993).
21. A. Sawano, S. Takayama, M. Matsuda, A. Miyawaki, *Dev. Cell* **3**, 245 (2002).
22. M. E. Vazquez, M. Nitz, J. Stehn, M. B. Yaffe, B. Imperiali, *J. Am. Chem. Soc.* **125**, 10150 (2003).
23. L. Guo *et al.*, *J. Am. Soc. Mass Spectrom.* **14**, 1022 (2003).
24. G. C. Brown, B. N. Kholodenko, *FEBS Lett.* **457**, 452 (1999).
25. B. N. Kholodenko, G. C. Brown, J. B. Hoek, *Biochem. J.* **350**, 901 (2000).
26. A. Nguyen, D. M. Rothman, J. Stehn, B. Imperiali, M. B. Yaffe, *Nat. Biotechnol.* **22**, 993 (2004).
27. P. Niethammer, P. I. Bastiaens, E. Karsenti, *Science* **303**, 1862 (2004).
28. The authors thank P. Verveer for the estimation of PTP1B kinetic parameters from fitting ES kinetics in vitro, E. Chuntharpursat and F. Haj for help with the experiments, and the anonymous *Science* referees for valuable suggestions. I.A.Y. was supported by Louis Jeantet Fondation de Médicine. The project was partially supported by grants R01 DK60838 and R37 49152 to B.G.N. and the European Union (LSHG-CT-2003-503259 to C.S. and LSHG-CT-2003-505520 to P.I.H.B.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/315/5808/115/DC1
Materials and Methods
SOM Text
Figs. S1 to S17
References

11 September 2006; accepted 21 November 2006
10.1126/science.1134966

Science

This abbreviated version of *Science*'s Information for Authors is printed in the first issue of each year. The current, complete Information for Authors is available at www.sciencemag.org/about/authors.

Science Contact Information

Phone: (1)-202-326-6550 (USA)
(44)-1223-326500 (UK)
Fax: (1)-202-289-7562 (USA)
(44)-1223-326501 (UK)
E-mail: science_editors@aaas.org
science@science-int.co.uk (Europe)

Science is a weekly peer-reviewed journal that publishes significant original scientific research, plus reviews and analyses of current research and science policy. Our main offices are in Washington, DC, and Cambridge, UK. We welcome submissions from all fields of science and from any source. Competition for space in *Science* is keen, and many papers are returned without in-depth review. Priority is given to papers that reveal novel concepts of broad interest.

CATEGORIES OF SIGNED PAPERS

Brevia (one page; about 800 words and one figure or table) are short peer-reviewed papers presenting novel results of broad general interest.

Research Articles (up to ~4500 words or ~5 journal pages) are expected to present a major advance. Research Articles include an abstract, an introduction, up to six figures or tables, sections with brief subheadings, and a maximum of about 40 references. Materials and methods should usually be included in supporting online material, which should also include other information needed to support the conclusions.

Reports (up to ~2500 words or ~3 journal pages) present important new research results of broad significance. Reports should include an abstract, an introductory paragraph, up to four figures or tables, and a maximum of about 30 references. Materials and methods should usually be included in supporting online material, which should also include other information needed to support the conclusions.

Technical Comments (up to 1000 words, with a maximum of 15 references) discuss papers published in *Science* within the previous 6 months. The authors of the original paper are given an opportunity to reply. Comments and replies are reviewed and edited as needed. Abstracts of the discussions appear in print; the full text appears online.

Reviews (4 journal pages, on average) describe new developments of interdisciplinary significance and highlight unresolved questions and future directions. All Reviews undergo peer review. They include an abstract, an introduction that outlines the main point, and brief subheadings. A maximum of 40 references is suggested. Most Reviews are solicited by the editors, but unsolicited submissions may also be considered.

Letters (~300 words) discuss material published in *Science* in the past 3 months or issues of general interest. They should be

Manuscript Preparation

See Information for Authors at www.sciencemag.org/about/authors for more detailed information.

One page in the journal contains approximately 1000 words and one small figure. Use double spacing throughout the text, tables, figure legends, and references and notes. Electronic files should be formatted for U.S. letter paper.

Titles should be no more than 96 characters (including spaces) for Reports, Research Articles, and Reviews, and 64 characters plus spaces for Brevia.

One-sentence summaries capturing the most important point should be submitted for all papers.

Abstracts explain to the general reader why the research was done and why the results are important. The abstract should convey the paper's main point and outline the results and conclusions.

Text starts with a brief introduction describing the paper's significance, which should be intelligible to readers in various disciplines. Technical terms should be defined. Symbols, abbreviations, and acronyms should be defined the first time they are used. All tables and figures should be cited in numerical order.

References and notes are numbered in the order in which they are cited, first through the text, then through the text of the references, and then through the figure and table legends. Each reference should have a unique number; do not combine references or embed references in notes. Do not use op. cit. or ibid.

Acknowledgments, including complete funding information, accession numbers, and any information related to authorship conflict of interest, should be gathered into the last numbered reference.

Tables should be included at the end of the references and should supplement, not duplicate, the text. The first sentence of the table legend should be a brief descriptive title. Every vertical column should have a heading, consisting of a title with the unit of measure in parentheses. Units should not change within a column.

Figure legends should be double-spaced in numerical order. The figure title should be given as the first line of the legend. No single legend should be longer than ~200 words. Nomenclature, abbreviations, symbols, and units used in a figure should match those used in the text.

Supporting online material (SOM) is posted permanently on *Science* Online, is linked to the manuscript, and is freely available. SOM includes materials and methods plus extra text, figures, tables, references, and video or audio clips that are important for the integrity of the paper.

Figures should be submitted as part of the online submission or, if necessary for large files only, on a CD. Allowable formats for submissions are PDF, PS, and EPS for illustrations or diagrams; and TIFF, JPG, PSD, EPS, or PDF for photography or microscopy. See our online Information for Authors for information on preparing art. Authors of accepted manuscripts will receive more specific information about electronic submission of art for publication. Do not send irreplaceable artwork. Most figures will not be relabeled by *Science* and will be printed at a width of 5.5 cm (2.25 inches or 1 column) or 12.0 cm (4.75 inches or 2 columns). Some illustrations (for example, bar graphs, simple line graphs, and gels) may be reduced to a smaller width. Symbols and lettering should be large enough to be legible after reduction. Avoid wide variation in type size within a single figure. In the printed version of the figure, letters should be about 7 points (2 mm) high. We can easily include high-resolution images as SOM.

Graphs should be labeled on the ordinate and abscissa with the parameter or variable being measured, the units of measure, and the scale. Scales with large or small numbers should be presented as powers of 10. Simple solid or open symbols reduce well. Avoid the use of light lines and screen shading. Instead, use black-and-white, hatched, and cross-hatched designs for emphasis. Use scale bars in place of, or in addition to, magnifications. When figures are assembled from multiple gels or micrographs, a line or space should indicate the border between two original images. No part of a figure may be selectively manipulated.

Units should be metric and follow SI conventions.

Submitting a Manuscript or Letter

Science now accepts submissions of manuscripts and letters only through our Web site: www.submit2science.org. We are not able to accept submissions by e-mail. Your Web submission should include a cover letter containing a statement of the paper's main point, any information needed to ensure a fair review process, and names of colleagues who have reviewed the paper for you.

Our online submission site cannot accept copies of related papers. They should be sent as a PDF by e-mail to science_editors@aaas.org along with your Web submission number. Very big figure files or movies may be too large for the Web submission site. These should be sent on a CD to *Science*, 1200 New York Avenue, NW, Washington, DC 20005, USA, or to *Science* International, Bateman House, 82-88 Hills Road, Cambridge, CB2 1LQ, UK.

submitted through our Web submission site (www.submit2science.org). Letters are not acknowledged upon receipt, nor are authors generally consulted before publication. Letters are subject to editing for clarity and space. E-letters are online-only, 400-word contributions for promoting rapid, timely discussions.

Policy Forums (1000 or up to 2000 words) present issues in science policy. **Education Forum** (2000 words) presents essays and original research on science education and its practice. **Books *et al.*** (up to 1000 words) present reviews of current books, multimedia, exhibitions, and films of interest to *Science* readers. **Perspectives** (up to 1000 words) analyze recent research developments but do not primarily discuss the author's own work. Most Book Reviews, Education and Policy Forums, and Perspectives are solicited by the editors, but unsolicited contributions will be considered.

MANUSCRIPT SELECTION

We are committed to the prompt evaluation and publication of submitted papers. Our submission and review process is fully electronic. Papers are assigned to a staff editor who has knowledge of the field discussed in the manuscript. Most submitted papers are rated for suitability by members of the Board of Reviewing Editors (see the masthead). The editors at *Science* consider this advice in selecting papers for in-depth review. Authors of papers that are not highly rated are notified promptly, by e-mail only, within about 1 to 2 weeks. Membership in AAAS is not a factor in selection.

Papers are reviewed in depth by at least two anonymous referees. Reviewers are contacted before being sent a paper and are asked to return comments within 1 to 2 weeks. We are able to expedite the review process for papers that require rapid assessment. Selected papers are edited to improve accuracy and clarity and to shorten, if necessary. Papers cannot be resubmitted over a disagreement on interest or relative merit. If a paper was rejected on the basis of serious reviewer error, resubmission may be considered. Some papers may need additional editorial oversight or present potential security concerns. Such papers will be brought to the attention of the Editor-in-Chief for further evaluation. If necessary, outside reviewers with expertise in the area will be consulted. Most papers are published 4 to 8 weeks after acceptance; selected papers are published online within 2 weeks of acceptance in *Science* Express (www.sciencexpress.org).*

SUBMISSION REQUIREMENTS

Authorship All authors must agree to be so listed and must have seen and approved the manuscript, its content, and its submission to *Science*. Any changes in authorship must be approved in writing by all the original authors. Submission of a paper that has not been approved by all authors will result in immediate rejection without appeal.

Prior publication *Science* will not consider any paper or component of a paper that has been published or is under consideration elsewhere. Distribution on the Internet may be considered prior publication and may compromise the originality of the paper. Please contact the editors with questions regarding this policy.

Human studies Informed consent must have been obtained for studies on humans after the nature and possible consequences of the studies were explained. All human research must have IRB approval.

Animal care Care of experimental animals must be in accordance with the authors' institutional guidelines.

Related papers Copies of papers submitted to other journals by any of the authors that relate to the paper submitted to *Science* must be included with the submission.

Unpublished data and personal communications Citations to unpublished data and personal communications cannot be used to support claims in the paper.

* *Press coverage and presentation at scientific conferences.* Reporting the main findings of a paper in the mass media may compromise the novelty of the work and thus its appropriateness for *Science*. Authors are free to present their data at scientific meetings but should not overtly seek media attention or give copies of the figures or data from their manuscript to any reporter, unless the reporter agrees to abide by *Science*'s press embargo. If a reporter attends an author's session at a meeting and writes a story based only on the presentation, such coverage will not affect *Science*'s consideration of the author's paper. The paper should remain a privileged document and should not be released to the press or the public before publication. Questions should be referred to the AAAS Office of Public Programs (202-326-6440).

CONDITIONS OF ACCEPTANCE

Funding and conflict of interest Authors must agree to disclose all affiliations, funding sources, and financial or management relationships, including those that could be perceived as potential sources of bias, as defined by *Science*'s conflict-of-interest policy, detailed at www.sciencemag.org/about/authors.

Data deposition Before publication, large data sets, including microarray data, protein or DNA sequences, and atomic coordinates and structure factors for macromolecular or chemical structures, must be deposited in an approved database, an accession number must be included in the published paper, and the deposited information must be released at the time of publication. Electron micrograph maps must also be deposited. Approved databases include GenBank or other members of the International Sequence Database Collaboration, worldwide Protein Data Bank, BioMed Res Bank, Electron Microscopy Data Bank (MSD-EBI), and SWISS-PROT, and as noted at www.sciencemag.org/about/authors.

License Authors retain copyright but must agree to grant to *Science* an exclusive license to publish the paper in print and online.

Data availability After publication, all data necessary to understand and assess the conclusions of the manuscript must be available to any reader of *Science*.

Access policies After publication, authors may post the accepted version of the paper on the author's personal Web site and are provided one referrer link that can be posted on a personal or institutional Web page, through which users can freely access the published paper on *Science*'s online site. *Science* allows deposition of accepted papers into PubMed Central 6 months after publication, in accord with the requirements of the funders NIH and Wellcome Trust, provided that a link to the final version published in *Science* is included. Original research papers are freely accessible with registration on *Science*'s Web site 12 months after publication.

Materials sharing After publication, all reasonable requests for materials must be fulfilled. Before acceptance, *Science* must be informed of any restrictions on sharing of materials [Material Transfer Agreements (MTAs), for example]. Unreasonable restrictions may preclude publication.

Motorized Stage System

The NanoScanZ Piezo Stage System is designed for researchers using deconvolution and three-dimensional imaging. The NanoScanZ features 100 mm or 200 mm travel, one nanometer repeatability, and closed-loop control using a subangstrom resolution Piezo resistive sensor. The capability for normal DIC (differential interference contrast) operation, quicker movement and settling time due to lower mass movement, and the added flexibility to create Z stacks with multiple objectives are major benefits of the system. The NanoScanZ complements the speed of the newest digital cameras by accomplishing in milliseconds what takes seconds for conventional rotary focus drives.

Prior Scientific For information 800-877-2234 www.prior.com

Glycoprotein Enrichment

Focus Glycoprotein rapidly fractionates glycoproteins that have terminal a-D-mannosyl and a-D glycosyl residues. Focus Glycoprotein makes use of spin columns containing the immobilized lectin concanavalin A for rapid glycoprotein isolation. The kit is supplied with all the necessary columns and buffers for 10 fractionations.

G-Biosciences/Genotech For information 800-628-7730 www.GBiosciences.com

RNA Purification

The MasterPure RNA Purification Kit can be used to purify high-quality RNA suitable for reverse-transcription and real-time polymerase chain reaction analysis from formalin-fixed, paraffin-embedded (FFPE) tissue. The quality and integrity of the purified FFPE RNA is excellent, with the 3′/5′ ratios of all tested specific gene transcripts being less than 2, according to the manufacturer. The total time from FFPE to pure RNA is less than two hours.

Epicentre Biotechnologies
For information 800-284-8474
www.EpiBio.com/masterpure_complete.asp

Proteomics Standard

The Universal Proteomics Standard is a complex mixture of 48 human proteins designed to enable researchers to better assess proteomics strategies, troubleshoot protocols, and normalize results from day to day and lab to lab. The protein mixture was evaluated by more than 100 independent proteomics laboratories worldwide.

Sigma-Aldrich For information 314-286-7616 www.sigma-aldrich.com

Microarray Dot Blot Kits

The Protein Detector Microarray Dot Blot Kits enable sensitive screening of biomarkers in high throughput format using antibody-based technology. They are suitable for use in the detection of low density reverse-phase protein microarrays and for profiling dozens or hundreds of samples for the presence of low abundant antigens. They provide significant sensitivity and specificity improvements over enzyme-labeled antibody methods through biotin-streptavidin amplification and require little assay optimization. Signal is further enhanced by the microarray blocker in the kit.

KPL For information 301-948-7755 www.kpl.com

Rapid Cell Separation

SpinSep is a novel separation method for human and mouse cells. It requires only a centrifuge, permitting multiple simultaneous separations. SpinSep is suitable for any sample, including frozen ones. Cells can be isolated from samples containing as few as 106 cells, with the maximum sample volume limited only by centrifuge rotor size. Unwanted cells are labeled with dense particles before centrifugation over a density separation medium. The unwanted cells pellet, while the desired cells are collected from the interface. High cell purities and recoveries can be achieved in approximately one hour.

StemCell Technologies For information 800-667-0322 www.stemcell.com

Protein Transfer Instrument

The iBlot Dry Blotting System is designed to transfer proteins from polyacrylamide gels to nitrocellulose membranes in preparation for downstream analysis in just seven minutes. Traditional wet or semi-dry systems for transferring proteins from gels to membranes rely on complex buffering systems and layers of filter paper, and can take several hours. The iBlot Drying System makes use of two solid-phase ion reservoir matrices in pre-assembled transfer stacks to blot proteins. The proprietary formulation of the transfer stacks and the integrated nitrocellulose membrane ensure reproducible protein transfers every time. The self-contained iBlot System does not require additional buffers or an external power supply, enabling western analysis to be completed in a single work day, rather than the 24 hours or more of traditional systems.

Invitrogen For information 800-955-6288
www.invitrogen.com/iblot

Peptide Synthesizer

The Activo-P11 Peptide Synthesizer is an easy-to-use and affordable instrument for synthesis of high-quality peptides. The automated and enclosed system performs a wide range of chemistries to synthesize peptides of any length in the 0.1–1.0 mmol scale.

Activotec For information +44-208-546-0869
www.activotec.com

For more information visit Product-Info, *Science*'s new online product index at *http://science.labvelocity.com*

From the pages of Product-Info, you can:
- Quickly find and request free information on products and services found in the pages of *Science*.
- Ask vendors to contact you with more information.
- Link directly to vendors' Web sites.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and government organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier by visiting www.science.labvelocity.com on the Web, where you can request that the information be sent to you by e-mail, fax, mail, or telephone.